

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 20:11:19 ; Search time 4424 Seconds  
(without alignments)  
10639.813 Million cell updates/sec

Title: US-09-993-777-1\_COPY\_9874\_10959

Perfect score: 1086

Sequence: 1 ATGGCATCGCTACTTGGAC.....GCCCGACTCAATAAAATG 1086

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.man.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1084.4	99.9	18912	14	ILU28832	U28832 Infectious
2	63	5.8	1627	14	HSMGDGLVCO	L31965 Gallid herp
3	51.2	4.7	125020	9	AF429315	AF429315 Homo sapi
C	47.8	4.4	2000	6	AX655393	AX655393 Sequence
5	47.4	4.4	163025	14	AY372243	AY372243 Peittacid
C	43.2	4.0	1246	6	AX164174	AX164174 Sequence
C	41.4	3.8	1598	8	AK119529	AK119529 Oryza sat
C	41.4	3.8	1618	8	AK065286	AK065286 Oryza sat
C	41.4	3.8	1756	8	AK121636	AK121636 Oryza sat
10	41.4	3.8	2000	6	AX655393	AX655393 Sequence
11	41.4	3.8	138025	8	AF003261	AF003261 Oryza sat
12	41.4	3.8	200183	8	AF003227	AF003227 Oryza sat
C	40.2	3.7	125020	9	AF429315	AF429315 Homo sapi
14	39.6	3.6	5513	3	AY135117	AY135117 Drosophil
15	39.6	3.6	5514	3	AY135123	AY135123 Drosophil
16	39.6	3.6	5521	3	AY135128	AY135128 Drosophil
17	39.6	3.6	5525	3	AY135131	AY135131 Drosophil
18	39.6	3.6	5526	3	AY135127	AY135127 Drosophil
19	39.6	3.6	19277	3	DMU13909	U19909 Drosophila
C	39	3.6	110000	2	LMFLCHR18_04	Continuation (5 of
21	38.8	3.6	507	8	AK062637	AK062637 Oryza sat
22	38.8	3.6	70000	8	AF003848	AF003848 Oryza sat
23	38.2	3.5	591	8	AK107720	AK107720 Oryza sat
C	38.2	3.5	16984	1	AE006954	AE006954 Mycobacte
C	38.2	3.5	40051	1	MSGY224	AD000004 Mycobacte
26	38.2	3.5	109555	8	CNS08C90	AL732849 Oryza sat
C	38.2	3.5	324050	1	EX248335	BX248335 Mycobacte
C	38.2	3.5	342416	1	EX842573	BX842573 Mycobacte
29	38	3.5	2677	6	AX777355	AX777355 Sequence
30	38	3.5	2838	6	AX777353	AX777353 Sequence
31	38	3.5	4516	3	BT001484	BT001484 Drosophil
32	38	3.5	4543	3	DROCSW	M94730 Drosophila
33	38	3.5	5492	3	AY135129	AY135129 Drosophil
34	38	3.5	5511	3	AY135134	AY135134 Drosophil
35	38	3.5	5521	3	AY135132	AY135132 Drosophil
36	38	3.5	5521	3	AY135133	AY135133 Drosophil
37	38	3.5	5522	3	AY135118	AY135118 Drosophil
38	38	3.5	5522	3	AY135119	AY135119 Drosophil
39	38	3.5	5522	3	AY135120	AY135120 Drosophil
40	38	3.5	5522	3	AY135121	AY135121 Drosophil
41	38	3.5	5522	3	AY135122	AY135122 Drosophil
42	38	3.5	5522	3	AY135124	AY135124 Drosophil
43	38	3.5	5522	3	AY135125	AY135125 Drosophil
44	38	3.5	5522	3	AY135126	AY135126 Drosophil
45	38	3.5	5523	3	AY135130	AY135130 Drosophil

# ALIGNMENTS

RESULT 1

ILU28832

LOCUS

DEFINITION

Infectious laryngotracheitis virus

glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein

E, ORF9 genes, complete cds.

U28832

U28832.1

GI:1486484

Gallid herpesvirus 1

Gallid herpesvirus 1

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.

1 (bases 1 to 18912)

REFERENCE

AUTHORS

Wild,M.A., Cook,S. and Cochran,M.

ILU28832 18912 bp DNA linear VRL 09-AUG-1996  
Infectious laryngotracheitis virus US10, US2, protein kinase, UL47,  
glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein  
E, ORF9 genes, complete cds.

TITLE	A genomic map of infectious laryngotracheitis virus and the sequence and organization of genes present in the unique short and flanking regions
JOURNAL	Virus Genes 12 (2), 107-116 (1996)
MEDLINE	97033380
PUBMED	8873127
REFERENCE	2 (bases 1 to 18912)
AUTHORS	Wild, M.A.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-1995) Martha A. Wild, Syntro Research Laboratory, 3535 General Atomics Court, San Diego, CA 92121, USA
FEATURES	Location/Qualifiers
source	1. .18912
	/organism="Gallid herpesvirus 1"
	/mol_type="genomic DNA"
	/strain="USDA challenge strain"
	/db_xref="taxon:10386"
repeat_region	<1. .2909
	/notes="large inverted repeat flanking the short unique region"
	/rpt_type="inverted"
	697. .1533
CDS	/codon_start=1
	/product="US10"
	/protein_id="AAC55093.1"
	/db_xref="GI:1486485"
	/translation="MENLDGCVPLAMDSHTAHVPRGRERRQGAVASSESASDR VPCIRASRLMRELVEISELKDVGSTESARDNELLANERLSAFAGASBART GLGLRPAWSTSVASPTDNNNGLEGELEAMGEGDFWLDSDGDFEDESRTM QSENMFVIEKELLSLRLPADLASAERTSRSLAAGHWCCLWHPRCREACLYD SIYVQSLFCVGTGRVQSEWRRREYLAALRAGAAAANSPEVSASIFARDGIALALAR RR"
CDS	complement(1900. .2784)
	/notes="coding region unique to avian herpesviruses"
	/codon_start=1
	/product="unknown"
	/protein_id="AAC55094.1"
	/db_xref="GI:1486486"
	/translation="MSKVCYCLARHLKSPRCVGRVAFGLATMSRPTSHLDLAFSA AFRGTLPGFRFRWASQSDIFFWPLAAVVOAKAYFEGKERLGSQVADITPAHD PRIAPAKAAVAAGALGTALGEGNGELESKWGKQIPRAAAWEIRDVPKVPVI GPDLISFFAAVELPVLVIRARGGASHRAHNNQSSAPAAGLAIRIGMENVRSLV IALPLNGLTPELDPGSGNSIRAFVAHLMNCVATDKIMSPDVRVPVEBSFYSHCLRE IINCEARFCVPCNPPKM"
misc_feature	2910. .16003
	/notes="short unique region"
CDS	complement(2916. .3605)
	/codon_start=1
	/product="US2"
	/protein_id="AAC55095.1"
	/db_xref="GI:1486487"
	/translation="MAPVKVTIIVSAVDSHYKLPNSRFELSDSGWKELVHAVKTMAS YD RPTSLVIVRPAVLVSGELFSLPRMCPVIRFGGDPGPGVPEWSGLDAGFYHLS SGAYAAKEPHLWLTADICMAALNPAPKTLITETGKNFERGVEIFLVNGDKTTL SLGSHVWTTTLARSSLRTPWFTVFLKVKFNSAAYCVSDSDGGERQPKFYLGLF SKXPSRRRR"
CDS	3694. .5124
	/codon_start=1
	/product="protein kinase"
	/protein_id="AAC55096.1"
	/db_xref="GI:1486488"
	/translation="MRFRICSRSAEKRRRTTENPLTSKRVCVLDSPFRTWLSRPYA EILPTAGVERLAELSVTTERAEVPTENTAVNSIPPANENGQNFAYAGDGFSTFK VDSHTDFDRASDYAGVPLAOTRLKHSDEFLOHFRVLDLVEGAYGFCIDVRVYTE BEGRRGVNSNQGSKCKELIAKYKNGTRASQLENILVIRLNHENVLKIQEL RYQNTVMYLRQFDLYSYNDEADFKDPSMLKTRIMKQKMSAVYIHSKILIH RDLKLENIFLNCQKTVLGFDTVPENEREPEFYGVGTVAINGFEILAKDSYCEI TDWSCVWLEWVSVEFCFPGCGGNHQQLKLVDSLSCDDEFPDPCNLNLYLH YASTDRAGHTVPSLRNLHLPADEVPLVQMLTFDWRLRPSAAEVLAMPLFSAEERT ITIIGHGKPIREIRARVPSNSEG"
CDS	5210. .7081
	/codon_start=1
	/product="UL47"
	/protein_id="AAC55097.1"
	/db_xref="GI:1486489"
	/translation="MTLPHRLTKRPARPCSVFVIHYSETKLDRYNKTMLLXRPDST MRHSGGDANHRGIRPRKSGIFASAREKTKGNALTESSSMDLDFSTKDFGKGW TVDGPADITAEVLQAWDLQVXKHEDAESEVVTYESKPTIQPFNAPDGPFWNAQD FTAPVYPSAEVLDAELVGVFVSRLQVQFTRSKSIVTRDAQSLGDSFMRIM QNVYTVLQCHITRLRHPSSKIVNCNDPLWYAYANQFHWGRVFLKASPPFYENRG QHGPMAVPRNAGAGLFLPAPWAAFEERDKLRLACLASLDLMDAAVLASRPYWRG VQDTSRFPALGCLSYFALVLLAETVLATMFEDHALVFWRALADGDFDYDTRIID PVNENLNGAEGTLLRGIIVASNTALVAVCANTYTIKLPVSATSNVAIRVETTLKA RRPQMDIYVRILOKPEFFFIAMQVATHANFCNLTKRSVDTGAPPLFRASSEKRL QOLNKMLCPLLVIOYEDFSKAMGSELKREKLETFVKATSSDRDPGRSLFLSDHAR EIIADGVRFKPVIDEFVRASVALSTAAAGVKARLLTSVRAPVPGAGAVSARKSEI"
	7245. .8123
	/codon_start=1
	/product="glycoprotein G"
	/protein_id="AAC55098.1"
	/db_xref="GI:1486490"
	/translation="MSGFNSIGSIATVSLVCSLLCASVLGAPVLDGLLESSPPFGGKI IAQACNRKTTIEVTPWSDYSGRTEGVSVKVFYGNPNPESFVGVDSESGHEDLS TCWALIHNLNASVCRSDAGIDPDFDKCEKVRRLRSGVELGSYVSGNGLVLYPGMY DAGIYAVQLSVGGKGYTGSVYLDVGNPNPGCHQGYTYYSLADENSLSSYDVASPEL DGPMEEDYSNCLDMPPLRPWTTCVSHDVEEQENATDELVLWDEECAGPLDEVYDERSE TMRMVVPSPPSTLQQ"
	8333. .11290
	/note="ORF5"
	/codon_start=1
	/protein_id="AAC55099.1"
	/db_xref="GI:1486491"
	/translation="MGTMVLRFLFLAVADAALPTGRCFCRVKVPVPGGTIOENLAVLA ESPVTHATVPPPEGAVSQIFADFTLRIYATDELBELAGTSSADNVTFSLS YRPRPEIHGAYFTIGVFATQSTESSYIGVSRVLVNASLERSVRLTTPCDENLQNPB TWGSKRWLGPSPVVRNDVAVLTAKAYIGECYSNSAAQTGLTSLMNTFFYSKRIYN VMTTGSPSPSRITVYSSRENGOPVLNRVSDGFLVKYPTDIDGRMINVANTSPADS GSULVATAPREGKLPSAIOLRHDMGTEPPGTTTDCOKMLETPLYALGNSVPRDD SIRGATLPFPDPAADFDGTSTPTTVEPAITLILPSTSDMGFFSTARGTSET LVSPQETDRLTTLPLTFPGESNTLFPPTAGISTEALQIAETELFTQSAEQMTT VFGQSTESSETARSQSEBPYFTQSTEQAAALQIAETELFTQSAEQMTT QPCABTEAQAQTPSTIPEIFTQSTQSTAPSAAPVFTQSSSTVTEVFTQPS TQPCATLASSTPAITRTTOSAGTEAFTQSSAEPDMRTQSTETHEFTQAPSTPKA TQPTSTPEVLITQSPSTPEVFTTILGAEPEITQPSAAPEVTVRSTSTMPETAOST LASQNTSSGSGTHNTEPRYTPVOTTHQKLTENKLTFFVTVRSTSEHMSAESOT PLDVKITLVYKFSNDGEPTATCVTSKVSPEVETNMKVDLVNDEISGNSPAGVFN S KNAQKQLYVYTDGTSVQMLCLSTSHSPPEYCLDFTSLIAREKDIAPELYFTSDP QYATCTILPSPVVRPEWSLNNVSLPEYLTATVTVSHTAGQSTVMSKSSAAGAWIS GRGNTIYECTVLISDGTVRTKERCITNTWIAVENGAAQAQLYSLFSLGSLGCGSI SALYATLWTAIYF"
	11098. .12402
	/codon_start=1
	/product="glycoprotein D"
	/protein_id="AAC55100.1"
	/db_xref="GI:1486492"
	/translation="MHRPHLRHSRYAKGEVLNKMDCGKRCSSGAAYFTLPWTCV RIMEHICFVRNAMDRLFLRNAPFTIVLLSSPASQSTAAVYDIILGRALDALITP AVGNRYLTVRSRGCVDVFNELPISNVDDMISAKEKKGKGPFEASVYVYIKGDDG EDKCYPIVRYKEGREGDVLQSCAVQAQWADVVPSTLVSRNAGAGLITFSPSTAAL SQYLLTLKGRPAQTALVTLVNDRLCKIGSOLNPLPSKWTTEQYTGQGEHLYP IADNTRHADDVYNGEYDILQNNLLRKNKPSAPDRPDSVPQEIIPAVTKAEGRTIP DAESSEKKAPEEDSDMQAEGENPAALPEDEVEDPEDEHDDSDPDYNDMPAV IPVEETTKSSNAVSMPIFAAFVACAVALVGLLVMSIVKCAR"
	12510. .13598
	/codon_start=1
	/product="glycoprotein I"
	/protein_id="AAC55101.1"
	/db_xref="GI:1486493"
	/translation="WASLLGTALLAALAPFGAMGVIITGNHVSARIDDDHIVIVAP RPEATIQLOLFFMPGQRPKPYSGTVAVFRSDITNOCYQBLSERENCTHRSVSYP VCKVYETTSASNELGPPPEKLTIRNRPNDSGMFEYVIRLDDTKEPDIYDPAIOL SVYQFANTATRGYSKASCTFGELPTVQLAEYLTERSHRNWQAYVATATTTTSAE TTPPTVATTSASELEASHFTFPLENGVDHTEPTPANENSVTVRLGMDFTLLGVTV AAVVSATIGLIVISIVTRNMCTPRKLDTVSQDDEBSQTRRESRFGPNVACEINK GADODSELVELVAIVNPSALSSPDSIKM"
	13792. .15291

```
/codon_start=1
/product="glycoprotein E"
/protein_id="AAC55102.1"
/db_xref="GI:1486494"
/translation="NMMLVILASCLARLTATRHVLFLEGTQAVLGBDDPRNVEPGT
VTKWKLNRNCKMAADVCSSPNYCFHDLIVDGGKDCPPAGPLSNALLIKRGES
FVVLGSLHNSNITNMTEYGGLLPDPVTRSDGIIYPRISQPDLANETTSYNVSL
SHVDEKAPHVEVDITIKPSAHARHVELQMLPFHLDNPSYTVTVPLRVFPPTFHV
KFNVTYSYGFVDFECEEVEKLFEPVYHPTDCKQFPATNQRCLIGSLVMAEFLGAA
SLLDGSRDLEBCHENRPNLRFDSRLSESAGLVISPLIAIPKVLITVWSDGILGW
SVTLGKRSRPNVVEITMPSKVPNKVIGSPGMDTGNVXYFVAGVAATCVIL
TCALLVKKKCAHQMGFSKTEPLYAPLPKNEFAGGILTDDEEVIYDEVYEPFRGY
CKQEPREDVNTFFGAVVEGERALNFKASALASWADRILANKSGRRNDSY"
CDS
15298..16080
/note="ORF9"
/codon_start=1
/protein_id="AAC55103.1"
/db_xref="GI:1486495"
/translation="MFFKTRGAEDAAAGKRNFKSRNREIILPRLGRTGKKTAGLSNY
TOPIPWNPFCARGESDNHAKDFTYRRTCCASSTVSSQDPDPSPTMPTEYGRVFS
AKRKLSDGCEGQPIVSKCLPDQSPARTYSQAORYTQVDSVSTPPGVDAVD
LETRAEPLGATTQOTESKNKLPNQSRLKPKPTNEHVGERCPSEGTVEAFSLGILSR
VGAATANELARWERTACPLAASAAAAGIVANMAARALQKQR"
repeat_region
16004..16912
/note="large inverted repeat flanking the short unique
region"
/rpt_type=inverted
16125..17013
/note="coding region unique to avian herpesviruses"
CDS
/codon_start=1
/product="unknown"
/protein_id="AAC55104.1"
/db_xref="GI:1486496"
/translation="MSKCYCLARHLKSPRCVGRVARGGLATMSRPTSHLDLAFSA
AFRGDLPGRWRASQCDIFFWPDLAIVQAAARVFEKERIGLSQVARDTAHD
PRIPAARAAVAAVGLMALSSELVGPNGELESKWGKQIPRAAAEIRDPVPI
GPDILFSAAVELPVLVYIRARGGAHSAHWNOSAPAAGLAIRIGMEVRSLLV
IALPLSNFTLPDLPEGGSQNSIRAFVAHLMNCVATDKIMSPDVRVPEESFYSHLRE
LIMCRAPCFPCNPPKKW"
complement(17380..18216)
/codon_start=1
Query Match 99.9%; Score 1084.4; DB 14; Length 18912;
Best Local Similarity 99.9%; Pred. No. 2.1e-290; Mismatches 1; Indels 0; Gaps 0;
Matches 1085; Conservative 0;
QY 1 ATGGCATCGCTACTTGGAACTCTGGCTCTCCTTGGCGGACGCTCGACCCCTTCGGCGCG 60
DB 12510 ATGGCATCGCTACTTGGAACTCTGGCTCTCCTTGGCGGACGCTCGACCCCTTCGGCGCG 12569
QY 61 ATGGGAATCGTGATCACTGGAATACACGCTCTCGCGCAGATTGACGAGATCATCGTG 120
DB 12570 ATGGGAATCGTGATCACTGGAATACACGCTCTCGCGCAGATTGACGAGATCATCGTG 12629
QY 121 ATCGTCGCCCTCGCCCCGAACTACAATTCAACTGCAGCTATTTTTCATGCTGCCAG 180
DB 12630 ATCGTCGCCCTCGCCCCGAACTACAATTCAACTGCAGCTATTTTTCATGCTGCCAG 12689
QY 181 AGACCCCAAAACCCCTACTCAGAACCGCTCGCGCTCGCTTGGCTGTGATATAACAAC 240
DB 12690 AGACCCCAAAACCCCTACTCAGAACCGCTCGCGCTCGCTTGGCTGTGATATAACAAC 12749
QY 241 CAGTGTCTACAGAACTTAGCGAGGAGCGCTTTGAAAATTGCACTCATCGTCGCTTCT 300
DB 12750 CAGTGTCTACAGAACTTAGCGAGGAGCGCTTTGAAAATTGCACTCATCGTCGCTTCT 12809
QY 301 GTTTTGTGCGCTGTAAGTACCGAGTACAGCTTCTCGGCTCGAACGACTACCGGA 360
DB 12810 GTTTTGTGCGCTGTAAGTACCGAGTACAGCTTCTCGGCTCGAACGACTACCGGA 12869
QY 361 CTTCCACACCGCTTAAAGTCACTATACGAATCTCTGTCGCAACGACGCGGGATGTC 420
DB 12870 CTTCCACACCGCTTAAAGTCACTATACGAATCTCTGTCGCAACGACGCGGGATGTC 12929
```

```
QY 421 TACGTAATTTGTTGGCTAGACGACCAAAAGAACCCATTGACGTTTCGGGATCCAACTA 480
DB 12930 TACGTAATTTGTTGGCTAGACGACCAAAAGAACCCATTGACGTTTCGGGATCCAACTA 12989
QY 481 TCGGTGTATCAATTCGGGAACCGCGGACCTCGGGACTCTATTCCAAAGGCTTCGTGT 540
DB 12990 TCGGTGTATCAATTCGGGAACCGCGGACCTCGGGACTCTATTCCAAAGGCTTCGTGT 13049
QY 541 CGCACCTTCGGATTTACCTACCGTCCCACTTTAGAGCCCTATCTCAGGACCGGAGGAAGTTGG 600
DB 13050 CGCACCTTCGGATTTACCTACCGTCCCACTTTAGAGCCCTATCTCAGGACCGGAGGAAGTTGG 13109
QY 601 CGCAACTGGCAAGCTACGTTGCCACGAGGCGCACGACACCGCGCGGAGGCGACAAC 660
DB 13110 CGCAACTGGCAAGCTACGTTGCCACGAGGCGCACGACACCGCGCGGAGGCGACAAC 13169
QY 661 CGGACGCGCTCACTGCAACACGAGGCTCCGAATTTGAAGCGGACACATTTTACCTTTCCC 720
DB 13170 CGGACGCGCTCACTGCAACACGAGGCTCCGAATTTGAAGCGGACACATTTTACCTTTCCC 13229
QY 721 TGGCTAGAAAATGCGGTGGATCATTAAGAACCGACACCCGCAACGAAAAATTCAAAAGTT 780
DB 13230 TGGCTAGAAAATGCGGTGGATCATTAAGAACCGACACCCGCAACGAAAAATTCAAAAGTT 13289
QY 781 ACTGTCGCTCTCGGACCAATGAGCCCTACGCTAAATTTGGGGTAAACCGTGGCTCGCTG 840
DB 13290 ACTGTCGCTCTCGGACCAATGAGCCCTACGCTAAATTTGGGGTAAACCGTGGCTCGCTG 13349
QY 841 AGGCAACGATTCGCGCTCGTCAATTTCAATTTCCATGTCACGAGAACATGTGACACCCCG 900
DB 13350 AGGCAACGATTCGCGCTCGTCAATTTCAATTTCCATGTCACGAGAACATGTGACACCCCG 13409
QY 901 CACCGAAAATTAGACACGCTCTCGAGACGAGAGAACGTTTCCCAAACTAGAGGGAA 960
DB 13410 CACCGAAAATTAGACACGCTCTCGAGACGAGAGAACGTTTCCCAAACTAGAGGGAA 13469
QY 961 TCGGAAAATTTGACCCATCGTTGCGTGTGGAATAAACAAGCGCGCTGACCAAGATAGT 1020
DB 13470 TCGGAAAATTTGACCCATCGTTGCGTGTGGAATAAACAAGCGCGCTGACCAAGATAGT 13529
QY 1021 GAATCTTGGAACTGGTTGCGATTGTTAACCCGCTTCGCTTAAGTCCGCCGACTCAATA 1080
DB 13530 GAATCTTGGAACTGGTTGCGATTGTTAACCCGCTTCGCTTAAGTCCGCCGACTCAATA 13589
QY 1081 AAAATG 1086
DB 13590 AAAATG 13595
```

```
RESULT 2
HSMGDGLYCO
LOCUS
DEFINITION Gallid herpesvirus 1 glycoprotein D (gp) gene, complete cds.
ACCESSION L31965
VERSION L31965.1 GI:493595
KEYWORDS glycoprotein D, Gallid herpesvirus 1
ORGANISM Gallid herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
REFERENCE 1 (bases 1 to 1627)
AUTHORS Johnson,M.A., Tyack,S.G., Prideaux,C.T., Kongsuwan,K. and Sheppard,M.
TITLE Sequence characteristics of a gene in infectious laryngotracheitis virus homologous to glycoprotein D of herpes simplex virus
JOURNAL DNA Seq. 5 (3), 191-194 (1995)
MEDLINE 95337426
PUBMED 7612933
FEATURES
source
Location/Qualifiers
1..1627
/organism="Gallid herpesvirus 1"
/mol_type="genomic DNA"
```

/strains="SA-2"  
/db\_xref="taxon:10386"  
/note="also known as infectious laryngotracheitis virus (ILT)"  
144..1517  
/gene="gd"  
144..151  
/gene="gd"  
326..1441  
/gene="gd"  
/codon\_start=1  
/product="glycoprotein D"  
/protein\_id="AAA98925.1"  
/db\_xref="GI:493596"  
/translation="MDRLHFLNFAFWITLVLSFASQSTAAVTVDYILGRALDALII  
PAVGYNYLTVSRGCELVNIPISNDVMIISAKEKEKGGPFASVVSFVYIKGDD  
GSGKYCPYIRKEKRGDQVLLSECAVQAQWALDYVSVTLVSRNGAGLTFIPNAA  
LGGQVLLTKIGRFACTALVLEVNDRLKIGSRNLNPLPSKWTTEQVQTFQGEHLI  
PIADNTNTHADVDYRVEDILORWNNLARKNPSAPDRPDVPOEIPAVTKAEGRT  
PDAESEKKAPEDESDMQAASGENFAALPEDDEVFEDTEHDDPNSDPPYNDMPA  
VIPVEETTKSNVAVSMPPIPAFVACAVASWGYWFGAS"  
1507..1517  
/gene="gd"  
polya\_signal  
ORIGIN  
Query Match 5.8%; Score 63; DB 14; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 3.7e-06;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCATCCTACTTGGAACTCTGGCTCTCTTGGCGGACGCTCGCACCTTCGGCGCG 60  
DB 1565 ATGGCATCCTACTTGGAACTCTGGCTCTCTTGGCGGACGCTCGCACCTTCGGCGCG 1624  
QY 61 ATG 63  
DB 1625 ATG 1627  
RESULT 3  
AF429315 125020 bp DNA linear PRI 18-JAN-2002  
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.  
DEFINITION  
ACCESSION AF429315  
VERSION AF429315.1 GI:17646244  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 125020)  
AUTHORS Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S.,  
Ingersoll-Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A.,  
Potter, N.T., Ross, C.A. and Margolis, R.L.  
TITLE A repeat expansion in the gene encoding junctophilin-3 is  
associated with Huntington disease-like 2  
Nat. Genet. 29 (4), 377-378 (2001)  
JOURNAL  
MEDLINE 21583737  
PubMed 11694876  
REFERENCE 2 (bases 1 to 125020)  
AUTHORS Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Margolis, R.L.  
TITLE Direct Submission  
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical  
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA  
FEATURES  
Location/Qualifiers  
1..125020  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16q24.3; between D16S520 and WI-12410"  
/note="isolated from a patient with Huntington's  
Disease-Like 2 (HDL2)"  
repeat\_region complement (35581..35746)

gene  
mRNA  
CDS  
/rpt type=tandem  
/rpt unit="ctg"  
complement(<36507..>36887)  
/gene="JPH3"  
/note="synonym: JP3"  
complement(<36507..>36887)  
/gene="JPH3"  
/product="junctophilin 3"  
complement(<36507..36887)  
/gene="JPH3"  
/note="component of the junctional complex between plasma  
membrane and endoplasmic reticulum"  
/codon\_start=1  
/product="junctophilin 3"  
/db\_xref="GI:17646245"  
translation="MSSGGKFNPDGSGYCGGWEDKHAHGVCTCPKQGGVTGWS  
HFEVLGVTPSGNTYQGTWAGKRGHGLGSKGWYKGEWTHGFKRGYGVRECA  
NGAKYEGTWSNGLQDGYGTETYSDG"  
ORIGIN  
Query Match 4.7%; Score 51.2; DB 9; Length 125020;  
Best Local Similarity 13.6%; Pred. No. 0.014;  
Matches 92; Conservative 275; Mismatches 305; Indels 4; Gaps 2;  
QY 47 CACCCCTCGCGCGATGGGATCGTGATCATCTGGAATCAGCTCTCGGCCAGGATTGACG 106  
DB 16839 CCCCCTGTGGACCCAGGTAGTAGCGGCCATCCARRKMKSHAGRRRCYYSKMSMSB 16898  
QY 107 ACATCATCATCTGATCGTCGCGCTCGCCCGCCCAAGCTCAATTAACCTCAGCTATTTT 166  
DB 16899 MSVSYVMSMSMASBSSCHWBKMTWSCMSMMYKSSWGWSCWCCWGRERRSKGWKY 16958  
QY 167 TCATGCTCGCCAGAGACCCCAACCCCTACTCAGAAACCGTCGCGCTCGCGCTTCGGT 226  
DB 16959 SRGMSRSKSMRYTGGSKMRRSSNMCTSSCYASCMCCWSSCCCMRSCCCRCYCCOMRY 17018  
QY 227 CTGATATAACAAACCACTGCTACAGGAACCTAGGAGGAGCGCTTTGAAAATTGCACTC 286  
DB 17019 CCACKCYMSYWTMSASYSMSRRSYWRSKMKWRCMSRRSCCKSRGCGSGMGKGGK 17078  
QY 287 ATCGATCGCTCTCTGTTTGTGCGCTGTAAAGTACCGAGTACAGTCTTCGCGCTCGA 346  
DB 17079 SYGRKTRKSRGKRGKAKWTMYRRSRMRKMTYSSKMGMYCWCWGRRCYCSMTSRS 17138  
QY 347 ACAGACTAACCGGACCTCCACACCCGTTTAAAGTCTACTATACGAAATCTCTCGTCCGAACG 406  
DB 17139 AMCCSYAKCKSNKCYSCYGMKGYIACSYRGSMSKCYMRGSTYSFSCGCCCTTTT 17198  
QY 407 ACAGCGGATGTTCTACGTAAT--TGTTCGGCTAGACGACCAAAAGAACCCATTGACGT 464  
DB 17199 TCCCCCNANTGGGAAGCTTTTNCNKYTSYRKRNGCAMCKYNNYNNYNSWRSSCRAGMS 17258  
QY 465 CTTCGCGATCAACTATCGGTGTATCAATTCGGGAACACCGCGCGACTCGCGACTCTA 524  
DB 17259 CTRYKSSMTMSASYSWMSMYTCMSRSMASRGSMSYMKMSWMSRMSYMKSCWCMK 17318  
QY 525 TTCCAAGCTTCGTGTCGACCTTCGATTAACCTACGTCCTCAACTTGAAGCCTATCTCAG 584  
DB 17319 MCYMMRSMRSGMSYMYASWSSSR--GCYCTRYCWNSSSKSYKSYMMRSMRSMK 17376  
QY 585 GACCGAGAAAGTGGCGCAACTGGCAAGCGTACGTTGCCACGAGGACCGACGACGAC 644  
DB 17377 MKWSRCWGSWGSASRSKCYKYSMRCSNMSSKCYRCAGCMWKMKGMYMRCWSMKRR 17436  
QY 645 CGCGAGGCGACAAACCCGACGCGCTACTGCAACACGAGCGCTTCGAACTTGAAGCGGA 704  
DB 17437 WVGKSAWYRVRWKGKGMWMMCMKCYSRMSRQWMMKMSYCASCRSCAMSGTYMKASY 17496  
QY 705 ACACCTTACCTTTCCC 720  
DB 17497 MCMSSYCTCCTCTSYM 17512



```
RESULT 4
AX655393/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Enarthroideae; Oryzaceae; Oryza.
1
REFERENCE
AUTHORS
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katahira, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE
Plant genes involved in defense against pathogens
JOURNAL
Patent: WO 03000898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source
1..2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 4.4%; Score 47.8; DB 6; Length 2000;
Best Local Similarity 7.5%; Pred. No. 0.065;
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;
QY 390 AATCTCTGTCGCAAGCAGCAGCGGATGTTCTACGTAAATGTTGCGGTAGACGACACCA 449
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
854 WMMWRITMYTYCYAMTCAKCKYKAMTKMTTWACAWRATSRWRAMAGMRKRYMKRA 795
QY 450 AGAACCATATGAGCTTTCGGATCCAACTATCGGTGATCAATTCGGGAAACACCGCCG 509
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
794 YMMWRWRCWAGWARKWMSRYKRWKWKYATRYKWKWMTWMSWRWMSYRWMSGMRM 735
QY 510 GACTCGGGACTTATTCAGAGCTTCGTGTCGACCTTCGGATACCTACCGTCCAAC 569
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
734 RWSAWRYCRRMKCAKTKYASSARFWTKRAKRSYRFRWYKRGWTKYRYRWSRMTTRA 675
QY 570 TGAGGCGTATCTCAGGACGAGGAAAGTTGGCGCACTGCGAAGCGTACGTCGACGGA 629
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
674 RMRKRKAGASWKSQWMTWRGARSWMYSKYSAKCKKTRMTSSYMSGTGMYMSY 615
QY 630 GGCACGACGACGAGCGCGGAGCAACCCCGAGCCGCTCACTGCAACGAGCGCCTC 689
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
614 KMSWTSKMSYMGKMTCTMTYSKMGSTRSKGWSGMSMYRWKWKYKRYMKW 555
QY 690 CGAACTTGAGCGGAAACATTTACCTTCCTGCGGTAGAAAATGGCGTGATCATTACA 749
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
554 KCTWRRCWYRMYGTYMTTTSRSMYMTGKRYKRYTSKRYMYKYRKYCWYTYGMYMKC 495
QY 750 ACCGACACCGGCAACGAAATTCACGTTACTGTCGCTCGGACATGAGCCCTAC 809
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
494 SYMMRYGYCKACCKCCYACWKAAYSGMMYMYTKYSKWRWMSMTKYMWSMYKKCRSMKY 435
QY 810 GCTAATGGGTAACGCTGCTCGCTGAGCGCAACGATCGGCGTCTGCTGATGTAAT 869
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 GAGCTGCKRWMTYCSYGMKWTYMSYKYSKYSKYSKYSKYSKYSKYSKYSKYSKYS 375
QY 870 TTCATCTGTCACGAGAAACATGTGACCCCGCAGCAACGAAATAGACACGCTGCG 924
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
374 YYYAKYKWKYKRRGTMYSYKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYS 320
RESULT 5
AY372243
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AY372243
Paistacid herpesvirus 1 isolate 97-0001, complete genome.
1
REFERENCE
AUTHORS
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katahira, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE
Plant genes involved in defense against pathogens
JOURNAL
Patent: WO 03000898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source
1..2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 4.4%; Score 47.8; DB 6; Length 2000;
Best Local Similarity 7.5%; Pred. No. 0.065;
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;
QY 390 AATCTCTGTCGCAAGCAGCAGCGGATGTTCTACGTAAATGTTGCGGTAGACGACACCA 449
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
854 WMMWRITMYTYCYAMTCAKCKYKAMTKMTTWACAWRATSRWRAMAGMRKRYMKRA 795
QY 450 AGAACCATATGAGCTTTCGGATCCAACTATCGGTGATCAATTCGGGAAACACCGCCG 509
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
794 YMMWRWRCWAGWARKWMSRYKRWKWKYATRYKWKWMTWMSWRWMSYRWMSGMRM 735
QY 510 GACTCGGGACTTATTCAGAGCTTCGTGTCGACCTTCGGATACCTACCGTCCAAC 569
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
734 RWSAWRYCRRMKCAKTKYASSARFWTKRAKRSYRFRWYKRGWTKYRYRWSRMTTRA 675
QY 570 TGAGGCGTATCTCAGGACGAGGAAAGTTGGCGCACTGCGAAGCGTACGTCGACGGA 629
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
674 RMRKRKAGASWKSQWMTWRGARSWMYSKYSAKCKKTRMTSSYMSGTGMYMSY 615
QY 630 GGCACGACGACGAGCGCGGAGCAACCCCGAGCCGCTCACTGCAACGAGCGCCTC 689
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
614 KMSWTSKMSYMGKMTCTMTYSKMGSTRSKGWSGMSMYRWKWKYKRYMKW 555
QY 690 CGAACTTGAGCGGAAACATTTACCTTCCTGCGGTAGAAAATGGCGTGATCATTACA 749
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
554 KCTWRRCWYRMYGTYMTTTSRSMYMTGKRYKRYTSKRYMYKYRKYCWYTYGMYMKC 495
QY 750 ACCGACACCGGCAACGAAATTCACGTTACTGTCGCTCGGACATGAGCCCTAC 809
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
494 SYMMRYGYCKACCKCCYACWKAAYSGMMYMYTKYSKWRWMSMTKYMWSMYKKCRSMKY 435
QY 810 GCTAATGGGTAACGCTGCTCGCTGAGCGCAACGATCGGCGTCTGCTGATGTAAT 869
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 GAGCTGCKRWMTYCSYGMKWTYMSYKYSKYSKYSKYSKYSKYSKYSKYSKYSKYS 375
QY 870 TTCATCTGTCACGAGAAACATGTGACCCCGCAGCAACGAAATAGACACGCTGCG 924
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
374 YYYAKYKWKYKRRGTMYSYKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYS 320
```

```
KEYWORDS
SOURCE
ORGANISM
Paistacid herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae.
1 (bases 1 to 163025)
Thureen,D.R., Keeler,C.L. Jr. and Dolan,M.
Pacheco's Disease Virus and Infectious Laryngotracheitis Virus:
Comparative Analysis of Two Avian Alpha Herpesviruses
Unpublished
2 (bases 1 to 163025)
Thureen,D.R., Keeler,C.L. Jr. and Dolan,M.
Direct Submission
Submitted (21-AUG-2003) Department of Animal and Food Science,
University of Delaware, 531 S. College Ave, Newark, DE 19717, USA
Location/Qualifiers
1..163025
/organism="Paistacid herpesvirus 1"
/mol_type="genomic DNA"
/isolate="97-0001"
/db_xref="taxon:50294"
complement(2376..4295)
/gene="UL54"
complement(2376..4295)
/gene="UL54"
/codon_start=1
/product="IE protein ICP27"
/protein_id="AA073686.1"
/db_xref="GI:34500203"
/translaton="MQABQTRCAAAGSAMESLHAAAGDEIPILHPPTPGAMSL
SDSLDTLAEIKEDLDLLGDPDIDDEEDQDFAELPPLYQPTENPPPNAA
NSTPEAQKOPAGKQNTERNERPMRKRKEDGVQPNPRRTHARSRSRAGSTSS
QOPSSSGARKPCVREAGDRETSKPKARGKYCAPPCAPACQOCSPACTASQKPP
PRQRPSTTAYSPDTPHQBHRHGGAGKMSVVTQDAKPPRRVPHRCGCTGR
HGRTKGRAGTSPVPLKPTREYKRNELKPKVMDIRSACTARDPARHDSG
AECPTKRDERTCAPQWRLREDNAAAFKDAAGLALPDDMCTNRDL
RAVADSVYTVARVSLYTNQCFPTMIGQSAAGVGMDSWELCRRAFLKAR
WSSQSVTVARVSLYTNQCFPTMIGQSAAGVGMDSWELCRRAFLKAR
LAQVQGLATPLGAVLHCHLAASKPLADPTSELPQATCLPTLFTADRFARW
HGHVTVLGHKVDHGVLTLYLPGMCAKIPVLDQHANVCRRECLLCAQLLG
KQTVGKFLCCDLA"
complement(5385..6461)
/gene="UL53"
complement(5385..6461)
/gene="UL53"
/codon_start=1
/product="glycoprotein X"
/protein_id="AA073687.1"
/db_xref="GI:34500204"
/translaton="MSRVOCLRLAAVIAISHLI FLVWFVCMNSVLENNEDCVVATRS
LAAQVQGLGAGNMTETRLIRGAAAPVSVGFNRSIYVIVNSASLLYSRETQD
GRCPANTFHKTDMAAVKMYVDGKNVVLVLEMDCAWILWFFQVRIATALLMLLAF
CVNRQRRGFGLDSASRVSAEAYLYNYTWTLAARVFLKRYLKLRFREIEYRREQ
TQFQSDTLGFLYMEPLALLRAIETILYFASLVSATVLRVNFDPSCVYLPNHVKV
FAMVFAALGALFVSAIDHLRETSRDAAAVIRPTNIIAACCANI ISHLLRLMY
GAALVIVGALKYREIQTLLG"
complement(6454..10335)
/gene="UL52"
complement(6454..10335)
/gene="UL52"
/codon_start=1
/product="DNA helicase primase component"
/protein_id="AA073688.1"
/db_xref="GI:34500205"
/translaton="MFGFLIGVFDRREESKCEPLVKHAPAIPEASEAENYPLSSA
DDKTLAALDAVETTLIGSCSLHVTDTAAWPAAGSAAAAMVMSGHSPODRLMIEAL
ESVRSALGTHSADTTCGQCSIALSAPQSPQSVPEGPINDSGQIFAVSGAGREENWI
DDEMGALCAADLQIDPIGDAQSVGGSTAACLPSPQKEEIRILITADGIPVVCISHL
GQALLQRYLYIAYSHVTPSAGGGLRATFYLLAATGRCRQQRQRPVPCVFTGGA
ADKTRALRGKPLWQDVIADEAQSALHEDLIALGTICAGTAATAAAGPCTGERN
LSYNYVASTPLGLRRTIASLMOYEVVAAFRKYNASFCPFWFKTFPGSETSLV
AVLYLLGATQDDGAGSDLQGLDVLTPFKPGNASGLTAARLNTAFERSFWC
CSELGAPKAEKIFREYDKTRVSSDTLLYQSLDSLIEKYRHALLPAXDFVRFYLAH
```

```
EGNRAAEAHMEALRNGGRABTAATRAQSBFFKGLREMCMPYFARNVDVAVKRL
PVRVAEHLRLSRYGTPRSGFWGFCWPSYDVSRLINKICAGPLCCLGTLQADAVRA
PPDGKGTCAETAATIIHSDSDASGESEATVKCGAHMPRANKRAGDGTIRISPAEC
ADTRAPLRLTAETAYTADGIMPTVLLYSLDTSAGDADAVKSEAKOORPVVY
RVAMPGGGAFAVARADOMKTVDVDFSGFWAPKSPESNGPLGSDKDAADDAWT
RALLARWGTSVARCAANVASOLYLNRNVLSDSLAVCNPLDLDITLKSSVGRP
SMULHMASVREAILVLLWSLFAADAEIDPTIYVIFYKTQCDAAAGDEEASPORRG
GGGDDVWANGEDVEIEEMGGYSQTAAMDMDWDADPQMDIDSDHPAEPMDD
GAGAAIBENKTAARAFCCRCRIGRIPCVPKPYALAGLSVAKATSLVQQAIVLQ
IARGIKCPKGTENLVLRQCTVFNKTSYSLVLYLSPGADTDGMDQVVAQTADR
TIMGDMVTLSHAMRVGDVDIPIHSODLAKMGLSAADSDTLDPIQLVGWMEPPSPPH
TADAREDAHVSPPERPOSSATLTSIDAKSRGLAQVPAS"
complement (11232..12476)
/gene="UL50"
complement (11232..12476)
/gene="UL50"
/codon_start=1
/product="dUTPase"
/protein_id="AAQ73690.1"
/db_xref="GI:34500207"
/translation="MEATTKQVAMIEIGKWAASCLARTCVISNEEDIYABRADTP
VLKDSGTIVSFSPGGVHARTVILKVPDDIMGACGAGARLPKTAITSFKGDEDLG
TSAGAGIVSFSPGGVHARTVILKVPDDIMGACGAGARLPKTAITSFKGDEDLG
NGRDHOMESLASIYPIILVQLDPCFYFCTGCKAFYRLGSLCTEPLNELGSHIY
LRGSAYESVFAADVFMVAMYGKWLIIQMAETPNKMLVELURDDSDSPAALIF
HDTFGKAEADAGYIRAPENCTLPFGSVRVILQKLMGKGRFAAFVNGRSMNLG
VLVEPVRVDDEWVSNTNIRDAAPFRKNDRILQVLALEDKLEIMGVDALPWRVV
QSVQBEKNSNRGDKFGSGGV"
12667..13014
/gene="UL49.5"
12667..13014
/gene="UL49.5"
/codon_start=1
/product="glycoprotein N"
/protein_id="AAQ73691.1"
/db_xref="GI:34500208"
/translations="MWLLRPAGSNFTVALIVLACAGPLTCSAQLDAGILNPNWGSAGHN
DAYMPGMFANSEDSRFPYSPHCSSRGLPLVNESMASVIFFLSLAMVCVAIVAILNCC
FNSFKNSVINSRW"
13184..14038
/gene="UL49"
13184..14038
/gene="UL49"
/codon_start=1
/product="vP22"
/protein_id="AAQ73692.1"
/db_xref="GI:34500209"
/translations="MSYVYNEQSGEKKTRKSSSRSDRKSASSSSPPPPGAVGRGV
MYPYGDDGAWLSPRRRSGSSSDSSSRDDDRRQPEQOPKQOSTRERKSTQTV
TYRKYTDGDDGGKSSNGSPWISDNOANLVKLSLAKGFSATPASPDSDRWRTTN
PWRSAFQVAVSVTMAQBELAARAWEKYKERNEDLEKIVETLEIKLITVNPGLW
DVASNTASAIRNQPIITHDLLYASSPAGDGARTSRQSCRKSMRPRGDEEDRAGSP
RPPSSRRR"
14500..16323
/gene="UL46"
14500..16323
/gene="UL46"
/codon_start=1
```

```
/product="vP11-12"
/protein_id="AAQ73693.1"
/db_xref="GI:34500210"
/translation="MSTKSSPARGTVVIPGHAMTLRSVELKKIPKILKTVFENAQKI
SPWLSRQASINSLAATAETMLPGILPSRAEHFNGDPTIPDKMVRVWVGSHKML
FRATLDTYHLYTRIEHREKFORARAMYTRASMLMDVELPFKDKMVGVAADL
EAICTHIVFWPHTVPAPNKSDDKAYLTSVAQFARSTAFIESYMYMPHPDEVNS
TDTNRVAAMLIIYHNGCKLASLFDALDAILRSKFNAGSRHLSGIYGFTRNGKA
TMAAKSGNVQIRRETCTCFPSDAMVECLISALDRMTYLVARWELMSVYKPGRTVP
IAGAGLAWLNGHLSYLTNMTVYVYLPHYLNPRLITISRRVLLQGLCTSHT
TERNATCPXMLLAQNAWEKQBERGTAALDVVCARAMRASRENKRPANITAMNNT
LEESAHVILKILBEAPOPAASAPSSLCATNYCPDSPSPRPVPOKRAKDPGRFSR
IMAGDDRRASVQSRGSGYWCHEHTASTKLRGRVLSAPLIHRRKSAEGLQGAAK
LRGSGRLPNGGRALPSPSPSSSESTDSPEARPETDHK"
complement (16521..17426)
/gene="UL45"
complement (16521..17426)
/gene="UL45"
/codon_start=1
/product="membrane protein"
/protein_id="AAQ73694.1"
/db_xref="GI:34500211"
/translation="MRNGARTQTAPPPAAEGTADVHNGEATITOLDGQPSAPSP
LTVFEATDDYKARNRNESEPEFVRAILDDDELSDIPVAAPLDVAPQKISCCGCVCL
RRCFVLTGMAVGLMAALATIAFAVPSVMTGVGVCERGWWQHVGVYKPGRTVP
PSSAADNGFAATPAGANTCAAAGSIVSDQALSFLLVLLSTASGSAAPRGSWM
TTHSGCAIVHYAPSSLLEVYNTSSLQSLALRTRDALDVSAIEDTCCSGSAGWCAAA
APAPTALKYAOLRAILTLGVRAE"
21070..22260
/note="ORFA"
/codon_start=1
/evidence=not experimental
/product="hypothetical protein"
/protein_id="AAQ73695.1"
/db_xref="GI:34500212"
/translation="MDLAGLSALAKYAVLASGLTOLVWTERDRISLMGPVLNPEG
LIPCTPFAVYDRIIRCPDRDLDMVRAHELTIGRTYWSLYSETFLNTEICIALP
LVPCRLGWKEAPPLASMERVASEVERLGDRTVNCVSRFAYIMGDMVITPEGVSI
RWDSGTGPFVFESEDACATVERVCRESFDIDSCVFPPLSTAPMTGNLFLTICLCVG
PRMTPAVRAQVEVLAKAIMQHTNSGRSFCNCPMETRDTPIGPVPSGSGIFIS
PEEKAEIARNGRLISLTKVVELVDGSRKPAVSWOSAWRKARVGLDPASVRLQ
TDAIDGSAAGRFLQOYCCTYVTRQLGFNERDARLRLRGGPVYVTTAPRGVVFDETT
DPLR"
22726..23829
/note="ORFB"
/codon_start=1
/evidence=not experimental
/product="hypothetical protein"
/protein_id="AAQ73696.1"
/db_xref="GI:34500213"
Query Match 4.4%; Score 47.4; DB 14; Length 163025;
Best Local Similarity 50.2%; Pred. No. 0.16;
Matches 145; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 156 CGAGTATTTTCATGCTGCGCAGAGACCCCAACACCTTACTCAGGAACCGTCCCGGT 215
Db 144244 CGCGCTCGTCTTTCTCGGAGACCGACCCCAAGACCCATACGGGGGACCGTAAGGT 144303
QY 216 CGCGTTTCGGTCTCATATAACAAACACAGTGTCTACAGAACTTAGCGAGGAGCGCTTGA 275
Db 144304 GCTTTTCAACCCCGGAGTCCGACGTGCTCTCATACCGTCTTCCAGGTACGACTCTC 144363
QY 276 AAATGCACTCATCGATCGTCTTCTGTTTTTGTTCGGCTGTAAAGTGACCGAGTACAGTT 335
Db 144364 TAACTGCACAAACAGCTCGCTCGGCTTCTTCGGCTGCTTACCGCACCGACCGAGTT 144423
QY 336 CTCGCGCTCGAACAGACTAACCGGACCTCCACACCGGTTTAAAGTCTACTATACGAATCC 395
Db 144424 CTCGCTGCGCGCGCAACACAGGGGTACGCTCTCCCGGCTTCG---TATCACTGAGGACCC 144480
QY 396 TCGTCCGACACAGCGGATGTTCTACGTAATGTTCCGGTACACGAC 444
```

Db 144481 CACCATGCTGACAGCGGCACATCATCTGCTGACCTCGATC 144529

RESULT 6

AX164174/c

LOCUS

DEFINITION Sequence 4 from Patent WO0138564.

ACCESSION AX164174

VERSION AX164174.1 GI:14545112

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Rouleau,G.A., Lafreniere,R.G., Rochefort,D., Cosssette,P. and Ragstate,D.

TITLE Loci for idiopathic generalized epilepsy, mutations thereof and method using same to assess, diagnose, prognose or treat epilepsy

JOURNAL Patent: WO 0138564-A 4 31-MAY-2001;

MCgill University (CA)

FEATURES

source

1. .1246

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 4.0%; Score 43.2; DB 6; Length 1246;

Best Local Similarity 12.7%; Pred. No. 1.2;

Matches 39; Conservative 141; Mismatches 127; Indels 0; Gaps 0;

QY 65 GAATGCTGATCACTGGAATACGCTCGCCGAGATTGACAGCATCATCTGTCGTCG 124

DB 445 SHEKSDSSSYMYKAAACNHNATMBCHBSABSCBSBCHGHBASDKYCNBTYKSY 386

QY 125 TCGCGCTCGCCGAGCTCAATCACTGCTATTTTCATGCTGCTGCGCAGAC 184

DB 385 SASNNYNYDYBESHYNSHSDHHTNSCHMTYCYSAVSNNYSSYSSDSSRYMAY 326

QY 185 CCACAAACCTTACTCAGAACCGCTCGGCTCGGCTTCGCTGATATACCAACCACT 244

DB 325 NCSYCMYSHSHHMCSSMYMYNYNYWTSNMSSTSSSHSYCTTSYSTATATTTT 266

QY 245 GCTACGAGCACTAGCAGGAGCGCTTTGAAATTCGACTCATCGATCGTCTCTGTTT 304

DB 265 MKTWTATNRKTBTBNTSCBKRACCTTYARNWHAKYSTWSAHSRNNYCTMBGRXC 206

QY 305 TTGTCGGCTGTAAAGTACCGGAGTACAGCTTCCTCGCTCGAACAGACTAACCGACCTC 364

DB 205 GCTHSSNCGTTCRDYSHRSMWBBANACNRBNANNSDSTNAGMNNYCKTBSGBAKHS 146

QY 365 CACACCC 371

DB 145 MMBSTCB 139

RESULT 7

AK119529/c

LOCUS

DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:001-203-C05, full insert sequence.

ACCESSION AK119529

VERSION AK119529.1 GI:37989152

KEYWORDS FLI cDNA; oligo capping.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1

AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team;

Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurotsaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsuura,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,T., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22753273

12869764

2

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurotsaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsuura,K., Matsuura,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Takami-Takeda,Y., Tagawa,A., Takahashi,F., Takaki-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

Unpublished

3 (bases 1 to 1598)

Kikuchi,S.

Direct Submission

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.

FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurotsaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsuura,K. and Murakami,K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurthara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES  
Location/Qualifiers  
1..1598  
/organism="oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="001-203-C05"

ORIGIN  
Query Match 3.8%; Score 41.4; DB 8; Length 1598;  
Best Local Similarity 53.4%; Pred. No. 3.8;  
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 560 CGCTCCACTTGGGCTATCTCAGGACCGAGGAAGTTGGCGCACTGGCGAAGGTACG 619  
Db 1424 CCATCCGGGTTACAGCTGCTGTTTGAAGATGCTGCTGGCGCGAGCGCGCATGT 1365  
QY 620 TTGCCACGAGGCGCAGCAGCAGCGCGGCGGACACCGCGCGCGCTCACTGCAA 679  
Db 1364 CGGCCACCGAGCGCCACCGCGGAGCAGCAGCAGCGCGCGCGCTTGGCGCT 1305  
QY 680 CCAGCGCGCTCCGAACCTTGAAGCGGAACCTTTACCTTTCCCTG 722  
Db 1304 GAGCGCGCCACCCACTCCGGCTGCTCTGCACTTCGCTG 1262

RESULT 8  
AK065286/c  
LOCUS AK065286 1618 bp mRNA linear PLN 24-JUL-2003  
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013002M14, full insert sequence.  
ACCESSION AK065286  
VERSION AK065286.1 GI:32975304  
KEYWORDS FLI CDNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1  
AUTHORS The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Sato, H., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Tsumoda, Y., Tida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsumoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Hara, A., Hashidume, W., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Kawai, J., Carninci, P., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kawaga, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
JOURNAL Science 301 (5631), 376-379 (2003)  
MEDLINE 22752273  
PUBMED 12863764  
REFERENCE 2 (bases 1 to 1618)  
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawayashi, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurthara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuki, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsumoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)  
Tel:81-29-838-7007, Fax:81-29-838-7007

COMMENT  
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cdna/  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsumoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Koya, S., Kurthara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES  
Location/Qualifiers  
1..1618  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013002M14"

ORIGIN  
Query Match 3.8%; Score 41.4; DB 8; Length 1618;  
Best Local Similarity 53.4%; Pred. No. 3.8;  
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 560 CGGTCCAACTTGGGCTATCTCAGGACCGAGGAAGTTGGCGCACTGGCGAAGGTACG 619  
Db 1485 CCATCCGGGTTACAGCTGCTGTTTGAAGATGCTGCTGGCGCGAGCGCTGCGCCATGT 1426  
QY 620 TTGCCACGAGGCGCAGCAGCAGCGCGGAGCGCAACCCCGACCGCTCACTGCAA 679

Db 1425 CGGCCACCGAGCCACCGCGGAGCAGCAGCAGCGCGCGCGGCGCTTGGCGCT 1366

Qy 680 CGAGCGCTCCGACCTTGAAGCGGACACTTACCTTCCCTG 722

Db 1365 GCAGCGCCACCCACTTCCGCGCTCCTCTGCACTTCCGCTG 1323

## RESULT 9

AK121636/c

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:J033051N12, full

insert sequence.

AK121636

AK121636.1 GI:37991259

VERSION

FLI CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of

Agricultural Sciences Rice Full-length cDNA Project Team;

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Nami, K., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

Ohtsuki, K., Shishiki, T., Foundation of Advancement of International

Science Genome Sequencing &amp; Analysis Group; Otsu, Y., Murakami, K.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Kie, Q., Lu, M.,

Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikeda, R., Ishibiki, J., Kawamura, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, Y., Ueda, M., Matsubara, K., RIKEN;

Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, T., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Ota, Y.,

Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,

Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,

Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamura, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,

Kodama, T., Kojima, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miura, J., Miyazaki, A.,

Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,

Nakamura, M., Nami, K., Nariakawa, R., Niikura, J., Nishi, K.,

Ooka, H., Otsu, N., Ota, Y., Otsu, Y., Otsu, R., Satoh, H., Sakai, C.,

Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,

Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,

Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,

Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Tanaka, T., Tomaru, A., Tota, T., Tsunoda, Y., Ueda, M., Waki, K.,

Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,

Yokomizo, S., and Yoshimura, A.

Collection, mapping, and annotation of 28K full-length cDNA clones

from japonica rice

Unpublished

REFERENCE

AUTHORS

Kikuchi, S.

Direct Submission

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of

Agricultural Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8502, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica  
rice.

URL : http://cdna01.dna.affrc.go.jp/cdna/

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Negata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, K.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,  
Yamamoto, M., and Nakahama, Y.

FATS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Iehibiki, J., Kawamura, M.,

Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M.,  
Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M.,

Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,  
Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N.,  
Ota, Y., Saitoh, H., Sakai, C., Sakai, K.,

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,

Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Tota, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1756

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J033051N12"

ORIGIN

Query Match 3.8%; Score 41.4; DB 8; Length 1756;

Best Local Similarity 53.4%; Pred. No. 3.9;

Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 560 CGGTCCAACTTGGAGCTCTCTCTGAGCCGAGGAAGTTCGGCAACTGGCAAGCTACG 619

Db 1485 CCATCCGGTTTCAGAGCTGGTTTGAAGATGTTGACGTGGCGAGCGCTGGCCATGT 1426

Qy 620 TTGCCAGGAGGCGCCAGCAGCGCGCGAGCGCAACCCGAGCCCGCTCACTGCAA 679

Db 1425 CGGCCACCGAGCCACCGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGCT 1366

Qy 680 CCAGCGCTCCGAACTTGAAGCGGAAACACTTTACCTTTCCCTG 722

Db 1365 GCAGCGCCACCCACTTCCGCGCTCCTCTGCACTTCCGCTG 1323

RESULT 10

AX655393

Sequence 5263 from Patent WO03000898.

AX655393

GI:29158207

AX655393.1

GI:29158207

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

AX655393

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriarthroideae; Oryzaceae; Oryza.

1

REFERENCE  
AUTHORS

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Nimura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Hajada, C., Hishigata, S., Honda, M., Ichikawa, Y., Idenuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T., Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakaishima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Sai, S., Sakai, K., Shibata, M., Shimokawa, T., Shumura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y., Yano, M., Jiang, J. and Gojobori, T.

THE genome sequence and structure of rice chromosome 1  
Nature 420 (6913), 312-316 (2002)

2237376  
12447438

REFERENCE  
AUTHORS

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission  
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
[E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468]

COMMENT

On Aug 9, 2002 this sequence version replaced gi:21202837.  
Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db/>) and the cDNA sequence database at RGP. Protein homologs of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.  
A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.  
The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0471B04 clone has an overlap with P0698A10 clone (DDBJ: AP003297) at the position 1 to 24,438 of 5' end and an overlap with P0018C10 clone (DDBJ: AP003227) at the position 48,801 to 138,025 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/genomeSeq.html>.  
Location/Qualifiers  
1. 138025  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clone="P0471B04"  
join(9533..9715,9887..10019,10850..10988,11365..11401,13640..13931,14823..14955,15564..15576)  
/gene="P0471B04.1"  
join(9533..9715,9887..10019,10850..10988,11365..11401,13640..13931,14823..14955,15564..15576)  
/gene="P0471B04.1"  
/note="hypothetical protein similar to Arabidopsis thaliana chromosome 1, F10C21\_14"  
/codon\_start=1  
/protein\_id="BAC07317.1"

gene

CDS

/db\_xref="GI:22202659"  
/translation="MSPPSVICQFGDTTYYKVFVGLAWETQKTRKRYFQFGEHLE  
AVITDNTGRSGKGFVTRFDGRACRACVDFPVIDGRNANCLNLYGVRSPPT  
POHGARSFRVMSQSQAQIGGLGAAPPSHATPHYAIQGLPHYHYGVSPPDY  
SYNTYNYFGAQTQFYGAATAAAGAAAGMTGSSPFVYFQFQSGSTTNTVASGQ  
GYNLOYPMFGSTVASTAAATGFAQQVGGFLSLAASPQAQAAEKANGKSNPNQAL  
TOGGFMILQANWKTKISQLIEQBCHFKISEK"  
complement(join(17199..17342,17914..18008,18141..18178,  
18784..19073))  
/gene="P0471B04.2"  
complement(join(17199..17342,17914..18008,18141..18178,  
18784..19073))  
/gene="P0471B04.2"  
/codon\_start=1  
/product="putative SnRk1-interacting protein 1"  
/protein\_id="BAC07318.1"  
/db\_xref="GI:22202660"  
/translation="WASGALARLLIGRRRAATPILARPPAKARASRRPQSPAPPS  
EDDFAGGEVAAAPTEGTSKLAELVRLGKRVPEALVKTRVEDGFSUKYIPWHLVVK  
NITLHAPESGVSIVSYSSDGKSVSYVYVTLGTDABIVREATGTSADDTGYGDP  
VOKAMAFRRACRTGLIGLHLHYHEDMS"  
join(24593..24925,25029..25403)  
/gene="P0471B04.3"  
join(24593..24925,25029..25403)  
/gene="P0471B04.3"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC07319.1"  
/db\_xref="GI:22202661"  
/translation="WARRQGVILRGRCRRRERGERQVRLPLASTVTVGRGGR  
GRRRGREGAGRTARQSSPRSSRRSSSSMTSRHQHILSDQRTGLSNAWDS  
TGLPVMSQVIVQRGAVRGRGRRGAGVAGDEGSGWRCGCHGRGRSRGA  
GMAAADGADAAAEADGLGGCGRRPIRRRAALLNQKSSSSRRSPSTPLAAPS  
PSAAVLLSGAMLPNR"  
join(25836..26033,26372..26471,26764..26924,27200..27397,  
27476..27568,27766..27888,28671..28918,29038..29209,  
29441..29527,29557..29820,30449..30611,30873..30960,  
31044..31113,31600..31728,32356..32445)  
/gene="P0471B04.4"  
join(25836..26033,26372..26471,26764..26924,27200..27397,  
27476..27568,27766..27888,28671..28918,29038..29209,  
29441..29527,29557..29820,30449..30611,30873..30960,  
31044..31113,31600..31728,32356..32445)  
/gene="P0471B04.4"  
/note="hypothetical protein  
similar to Arabidopsis thaliana chromosome 5, At5g49880"  
/codon\_start=1  
/protein\_id="BAC07320.1"  
/db\_xref="GI:22202662"  
/translation="WILRTPQKRRADSDVDAATAATTRSPVSDRLVLYDRPTAL  
VFAGVPFPMDDMCTYHCRQMVSEFVALDTAEKQVQYRATLDDMEERLSKSEDE  
RATCDKLNLYEQLAATKGRSAQERLLKEVGDFQRYCDQIKKIGELSTQLKEI  
DSRIAESSASAKESVLEGNLQRLSENSEKALKKLSYLQDDTKLSISKLANA  
ELERMLRALNSEDRAKLNLEOLDKQDESRENNEMEHRLNCSLSYERTPSD  
DQKLIQLLEELRYNEKEVDEARRLKSHTNVLLKEXLIEQCGRERAEMLSKQKE  
ISAKQKLEELNACSTALLUSNPVSSFGDIPQKIADIKQKALTNLNKGVSITQKE  
LKVALEFADLSQREAGEATLAKREASATREIKRELILAAISEERDRKRAVSK  
SRGDADSSQCFATILNGMNIISCEVTIEMESLSRMEKVTVELESTIRDORELIS  
COHTENLNMEXLSTESRKAKSLERDQOLRSQVALLSKLGHGYSASSTKVLRMVN  
TLAVNEAKQTLLEALQAKTKERLQAVEELKQADVTGVVDVNIAEKLAQKNQIA  
TEKREERYKAVPARIISVFRKACCSLFYKILVMDQQQNGIPVTRILHSVVAQSD  
DEKLEFDYEGSTNIIVDFIRMNISIPAFNTLMTESFNKRSIC"  
join(33470..33475,33573..33794,33969..34033,34129..34322,  
34573..34839,34942..35029,35125..35250,35325..35434,  
35557..35705)  
/gene="P0471B04.5"  
join(33470..33475,33573..33794,33969..34033,34129..34322,  
34573..34839,34942..35029,35125..35250,35325..35434,  
35557..35705)  
/gene="P0471B04.5"  
/note="contains ESTs C91866(E31978), AUI66151(E31978)"  
/codon\_start=1  
/product="putative elicitor response protein"

/protein\_id="BAC07321.1"  
/db\_xref="GI:22202663"  
/translation="MMTLSTSERQPOEKPRTRPPLPKAVAALCVAFVVGILLIS  
GRVVVLLPPGSSPSASNYKTSFTGCENRAKLGESNPTDINNEVSRTHAIQSLDK  
AVSLEMLAVERARSAACAGTAAAGLPQKAFVIGINTAFSSKKRDSLRDITVW  
PRDKLRLKEKGIIVRFVIGRSGAAAGDPLDRAVDAEDAKDFLRLDHVEGYH  
ELSKRTVYTTAVATWADAPYKYVDDDVHVNLMGLTSRLAKYTRPRVYVGVKSGCP  
VLQKGVKYEHEPVRKPFDEGNKYFRHATCGIYAVSADLAAYISINQPIILHRFANEVH  
SLGAWLGLVEHVRDMSMCCATPPDCEWKKRAGNVCVASFDFWSCSGCVKSDVRMKHI  
HRACGEGQAGVSWAT"  
join(36438..36570,36637..36956)  
/gene="P0471B04.6"  
join(36438..36570,36637..36956)  
/gene="P0471B04.6"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC07322.1"  
/db\_xref="GI:22202664"  
/translation="MSLTFESLVVPSRGRPRPQVRGRGSSSSVGGGGBESHWTVG  
TAEPRERDRRRRRRRRVTLPPNNIGACPLSPRCGADGTAFTSRSGGPDAP  
TPADPLRPTVTRDGEVSDGARQQRQWMDGQRGRKEREREAATADK"  
complement(join(38223..38304,39187..39287))  
/gene="P0471B04.7"  
complement(join(38223..38304,39187..39287))  
/gene="P0471B04.7"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC07323.1"  
/db\_xref="GI:22202665"  
/translation="MVKNDRKSVTGLTPTVCDWTDNNAPTSHLLPAEWSTFSHSIDDT  
CKVIRKIIIVSYSAIG"  
join(43057..43393,43868..43997,44238..44453,44554..45171,  
45260..46109)  
/gene="P0471B04.8"  
join(43057..43393,43868..43997,44238..44453,44554..45171,  
45260..46109)  
/gene="P0471B04.8"  
/note="contains ESTs AU057294(S21318), AU057295(S21318)"  
/codon\_start=1  
/product="putative male sterility 1 protein"  
/protein\_id="BAC07324.1"  
/db\_xref="GI:22202666"  
/translation="MVVNSRPLKRRTRVEARDFAGPPAGDGAAGTFREAVRGFLA  
RYARLLPLPSIFSPAAAAAPPHTLWVSLRVGEEGDEGGGAGVENVVEEDVLRSR  
SVYCDQCRVVGNSGHPVCGKRYHFIENNNDNQVCGKRHSCLCGTPTVAGESRCLLC  
NFDMDGELEECGMHLDNTHLHVAHVHANGHLLRVNREGGRSCLTGRDIMSFW  
Query Match 3.8%; Score 41.4; DB 8; Length 138025;  
Best Local Similarity 53.4%; Pred. No. 7.5;  
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
Qy 560 CCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAATTTGGCGCAACTGGCAAGCGTACG 619  
Db 95038 CCATCCGGTTTCAGAGCTGCGTTTGAAGATGCTGAGCTGCGCAGCGCTGCGCCATGT 95097  
Qy 620 TTGCCACGAGGCGCACGACGACGCGCGGAGGAGCAACCCGACCGCTCTACTCAA 679  
Db 95098 CGGCGACCGAGCCACCGCGCGAGCGACGACGACGAGCGCGCGCGCGCTTGGCGCCT 95157  
Qy 680 CCAGCGCTCCGACTTGAAGCGAACACATTACCTTTCCTG 722  
Db 95158 GCAGCGCACCCACTTCGCGCTGCTCTCTGCACTTCGCGCTG 95200  
RESULT 12  
AP003227  
LOCUS  
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
PAC clone: P0018C10.  
ACCESSION AP003227 BA000010  
VERSION AP003227.3 GI:21952783  
KEYWORDS  
SOURCE Oryza sativa (japonica cultivar-group)



ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzoideae; Oryza.

REFERENCE  
AUTHORS  
1  
Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,  
Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y.,  
Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,  
Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,  
Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,  
Idonuma, A., Iijima, M., Ikeda, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T.,  
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,  
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,  
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaka, H., Nakashima, M.,  
Nakana, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,  
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,  
Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K.,  
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,  
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y.,  
Yano, M., Jiang, J. and Gojobori, T.  
The genome sequence and structure of rice chromosome 1  
Nature 420 (6913), 312-316 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
22373736  
12447438

REFERENCE  
AUTHORS  
2 (bases 1 to 200183)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kannondai  
1-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:htp://rgp.dna.affrc.go.jp/,  
Tel:81-298-38-7441, Fax:81-298-38-7468)

JOURNAL  
On Jul 24, 2002 this sequence version replaced gi:21202836.  
Genes were predicted from the integrated results of the following:  
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor  
(October 1998 version). The genomic sequence was searched against  
NCBI NonRedundant Protein database, nr  
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
RGP. Protein homologues of the coding regions were searched against  
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent  
the identified cDNA sequences using BLASTN 2.0 with the  
corresponding DBJ accession no. and RGP clone ID.  
A gene with identity or significant homology to a protein is  
classified based on the 'putative-' and '-like protein'. A gene without  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with EST homology (covering  
almost the entire length of partial sequence) is classified as an  
'unknown' protein. A gene predicted with a gene prediction program  
is classified as a 'hypothetical' protein.  
The orientation of the sequence is from SP6 to T7 of the PAC clone.  
This sequence of P0018C10 clone has an overlap with P0471804  
clone (DBJ: AP003261) at the position 1 to 89,225 of 5' end and an  
overlap with B1065810 clone (DBJ: AP003561) at the position  
158,410 to 200,183 of 3' end. Detailed information on overlap  
assembly quality together with annotation of this entry is  
available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES  
source  
1. 200183  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clones="P0018C10"  
join(6052..6159,6257..6388,6483..6612,6719..6805,  
7232..7302)  
/gene="P0018C10.1"  
join(6052..6159,6257..6388,6483..6612,6719..6805,  
7232..7302)  
/gene="P0018C10.1"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC06200.1"

/db\_xref="GI:21952784"  
/translation="MWTSLIARSTTCLQTANDTTAPAKPAGAGPOLKLLADADISLE  
EFGSLCPPICTGELLITPDSNFAVVTRELLYQVTRLCRCGFGVSTOICHNLVDA  
AGITQFLQAVGELVEGAERPADELVHRRFLSGDFDVAALRDQLPRLSPGCHVSKTK  
QKYCQETFGSEWFERV"  
complement(join(14216..14383,14464..14772))  
/gene="P0018C10.2"  
complement(join(14216..14383,14464..14772))  
/gene="P0018C10.2"  
/note="contains EST AU058082(B30867)"  
unknown protein  
/codon\_start=1  
/protein\_id="BAC06201.1"  
/db\_xref="GI:21952785"  
/translation="MGRSRPWFQDPSAPPQVPPVAAAADAGEASTSLRDTGTSMD  
AISFGFAATAILVSMFLMIAFEHLKPHVFPPLASALPARRRHGVSPAGKLRSP  
MVETVQAADLSVLMPCQRYPTTAPAPLPAPCPREBHVHPDHDVHHSYMP"  
complement(join(21230..21448,21542..21620,22147..22244))  
/gene="P0018C10.3"  
complement(join(21230..21448,21542..21620,22147..22244))  
/gene="P0018C10.3"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC06202.1"  
/db\_xref="GI:21952786"  
/translation="MGYGRILSLRSLVTSLNQLNPTDNLHQNNTKHADVETDDLP  
KNKGALSNVSAEKVEAKKANKKRRRSGYSIRPPVHGHAASDRSSNPLGDPL  
DFLLTMGERAVHLPAKRPVPAKASIRPSP"  
join(26342..28883,29797..30149)  
/gene="P0018C10.4"  
join(26342..28883,29797..30149)  
/note="contains ESTs"  
C96750(C10643),C22379(C12740),C22380(C12740)"  
/codon\_start=1  
/product="putative receptor protein kinase"  
/protein\_id="BAC06203.1"  
/db\_xref="GI:21952787"  
/translation="MLPQQLQIYLCFILLSLKFGISASLPLETALLDICKSHLEDDPN  
YLGNDSSHSPCQYGVTCQTSQGVIGISLNASLSGTSSSFSLQSRLTELGAN  
SISTIPAAALNCTNLOVLNLSLTGKQLPDLSTFINLOVLDTNNFSGFPFWG  
KSLGTTELGLGNNFNEGDYPSIGLKLNTLWFLGQCNLRGELPVSIFDLVSLGTLID  
FSRQIIGVPIALSNRLNWKILYQNLITGELPPELAHLLTLLSEPDVSNQSLGIL  
PEKIANLKQKIFHYRNNSGVLPGLDLEFLESFTYENQSGKFPANLQRPSPL  
NAIDSNYSFGEPRFLCQNLQFLLDNNFSGFPSSYCKTLQRFRIQNLQ  
TGRTHGWLPNAVIVDANKFVGISDGIQSASLNOLYVHNYSFSGELPMLNLA  
LSLQKLVAFNNRSGQIPAQIGLSKOLSLHLEQNALGSIIPDIDGMCNSLDVLA  
DNSTGIPIDTLASLFTLSNLNLSHNMISGEIPEGLQYKLSYVDFSHNNLSGPVPA  
LLMTAGDDAFSENDGLCIAGVSEGRQNAINLYCPWNDHQNFSQRLLFVLLIVTS  
LVLLSLGLCLRYENYKLEQFHSKGDIESGDDSDSKVLESFHPPELDPESICLDVD  
NLICGGTGKVVYRLKSGRGAVALKQWRDDAKVMRTINTLTKIRHRLILKHAF  
LTGSESLVYVYVNGLYDAIRREFKAGQPELDWERYIAVGTAKGMYLHDCS  
PAIHRDIKSTNILLDEYAKLADFGIAKLVEGSLSPCAGTHGYNAPELAYSLSKT  
EKSDYVSGVILLLETSRPSQDQDGLDIVSWSSHANQNPAAVLDPKVSASHAS  
EDMTKVNIAILCTVQLPSPRTWRVYVLMIDISANGAKNNKDKK"  
complement(join(30791..31219,31321..31545,31728..32088,  
32199..32292,32370..32603,33125..33179))  
/gene="P0018C10.5"  
complement(join(30791..31219,31321..31545,31728..32088,  
32199..32292,32370..32603,33125..33179))  
/gene="P0018C10.5"  
/codon\_start=1  
/product="putative amino acid permease 6"  
/protein\_id="BAC06204.1"  
/db\_xref="GI:21952788"  
/translation="MNKNAAPEDESSEHRTGTWTATAHIIVATIGSGVLANSVS  
AQLGWAGPLALAGFACVYTTISWVAIRSDCIHRNGAGAAACVATVLMFLAF  
MCGTAQVYVNLGTVGTITATISWVAIRSDCIHRNGAGAAACVATVLMFLAF  
VQVLSQFPGLHITWYVAVMSFAYSFIGLSVAEVAWSHGSLSGRIQAGATA  
SSSKIMNVLLALGNITAFVAVLIEIQTLPSPENKMKKAAWYGATTFY  
ISVCGAYAAFGSDPAGNIPASGFPFLVDIANMCLILHLIGIYQVADIPATME  
RWISRWPEAKFINSVTVNPLIQSGSVTVAPYKLVLTVTVATTVAWMIFFPFA  
VLGLLGAFSFWPLTVYFPISMHIAQEKITRGRWYLLQGLSMVCLMISVAVIGSVTD





```
Db 71347 GACCATCT 71339

RESULT 14
AY135117
LOCUS
DEFINITION
Drosophila melanogaster isolate 1 corkscrew phosphatase splice
variant (csw) gene, alternatively spliced, complete cds.
ACCESSION
AY135117
VERSION
AY135117.1 GI:23344935
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 5513)
Riley, R.M., Jin, W. and Gibson, G.
AUTHORS
Contrasting selection pressures on components of the Ras-mediated
signal transduction pathway in Drosophila
TITLE
Mol. Ecol. 12 (5), 1315-1323 (2003)
JOURNAL
MEDLINE
22582092
PUBMED
12694293
REFERENCE
2 (bases 1 to 5513)
Riley, R.M., Jin, W. and Gibson, G.
AUTHORS
Direct Submission
TITLE
Submitted (23-JUL-2002) Department of Genetics, North Carolina
State University, Gardner Hall 7614, Raleigh, NC 27695, USA
JOURNAL
FEATURES
source
1. .5513
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/isolate="1"
/db_xref="taxon:7227"
/country="USA: Ann Arbor, MI"
/note="genotype: isofemale"
<161..>5449
/gene="csw"
join(<161..174,757..944,1017..1143,2096..2258,2951..3130,
3199..3316,3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/product="corkscrew phosphatase splice variant A"
join(161..174,757..944,1017..1143,2096..2258,2951..3130,
3199..3316,3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/note="CswB"
/codon_start=1
/product="corkscrew phosphatase splice variant A"
/protein_id="AA117607.1"
/db_xref="GI:23344936"
/translation="MSSRRWFPHPTISGIEAEKLLQEQFGDGSFLARLSSNPGFTLS
VRGNEVTHIKIONGDPPDLYGKGFATLPBVQYWMENGELKKNQOATLKLQPLI
CAEPTTERFHNLGSKAEKILIRKNGSFLVRESQKPGDFVLVSRDQKVTVM
IRWQKIVDGGESFGTSELIDYKKNPMVETCTVHLRQPNFNAITAAGINAR
VQLVKGVEFEFESLQDSRTFRNKGYSKQENLKNRYNLPDTRVRLKLDVEH
SVAGAEINANTIRLPTDGLYMSSESLSNSVSPCPACTAQTQRCNSCLOQK
TCVQCAVKAILPYNCACTCRKSDLSLKHKSESASSPSGSGSGSGTSGVS
SVNGPCTNLSTAGTCLVGLKRNDSGAVISMAEREREREMFXYATQG
CLLTQVNTDPMWQENTRVIMTKYERKEKCAVWPDGSEQFCHARIO
CVSNTSYTLRFLVSRDQPARFIHFQVWPDGHPADPCVLNPLQVNTRO
SHLAQERKPGICVHCASGIGTGTIVIMDLQIVRNGDITDITQRTQVRSQ
RSLGVEAQYKVFYAVQVHYQTLIAKRAEQSLQVREYTNKYTGIEGNSORS
PLPAPTSSILVSPKLTPTSDTLGTSGLSMGVGMGNKHSKQPPPLPVNCCN
NNGNIGNSCSGSGSTTSSNGSNGINALGGIGLGGNNRKNFYSLSLQK
QOREQAPAGAKWQFAPLPRPGILKLTSPVIFQNSKTFPKT"
join(<1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/product="corkscrew phosphatase splice variant B"
join(1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"

gene
mRNA
CDS
FEATURES
source
Query Match 3.6%; Score 39.6; DB 3; Length 5513;
Best Local Similarity 52.4%; Pred. No. 15;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 569 TTGAGGCTATCTCAGACCGAGGAAAGTTGGCGCACTGGCAAGCGTAGTTGCCACGG 628
Db 3979 TTAAGACCTACATCGCCACCCAGGGGTGTCTGTCTACCCAGCAAGTGAACCGGTACGG 4038
Qy 629 AGGCACGACGACGAGCGCGGCGGACCAACCCCGCGCGCTCACTCAACACCGCGCCT 688
Db 4039 ACTTTGGAACATGGTCTGGCAGAGAACACGCGGGTGATCGTCATGACACCAAGAGT 4098
Qy 689 CCGAAGCTTGAAGCGGACACCTTTACCTTCCCTCCCTGGGTAGAAAATGG 734
Db 4099 ACGAGCGCGCGCAAGAAAGTGGCGCGCTACTTGGCCAGCAGAGGG 4144
RESULT 15
AY135123
LOCUS
DEFINITION
Drosophila melanogaster isolate Reids2 corkscrew phosphatase splice
variant (csw) gene, alternatively spliced, complete cds.
ACCESSION
AY135123
VERSION
AY135123.1 GI:23344953
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 5514)
Riley, R.M., Jin, W. and Gibson, G.
AUTHORS
Contrasting selection pressures on components of the Ras-mediated
signal transduction pathway in Drosophila
TITLE
Mol. Ecol. 12 (5), 1315-1323 (2003)
JOURNAL
MEDLINE
22582092
PUBMED
12694293
REFERENCE
2 (bases 1 to 5514)
Riley, R.M., Jin, W. and Gibson, G.
AUTHORS
Direct Submission
TITLE
Submitted (23-JUL-2002) Department of Genetics, North Carolina
State University, Gardner Hall 7614, Raleigh, NC 27695, USA
JOURNAL
FEATURES
source
1. .5514
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/isolate="Reids2"
/db_xref="taxon:7227"
/country="Portugal: Mediera"
/note="CswB"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
/protein_id="AA117608.1"
/db_xref="GI:23344937"
/translation="MLFNKLEKLSLSSIGNVNVHKLQEQVYNNNNNNNNNNNNNNNNNN
NAYNORNEFEYERAIQAHVGSKRSEERESGKFKASKGRYAKVTPPTETPAQBP
CKNCWTHDELAQILKGVAGDAQRNRRLQRRRPLSAQPSAASASTSTESLHRL
TPSPQASYPATPTSTATPPQAPAFGASCSNSTLSLATMRVQLHGTWTHGNJUSG
KEAKLILRKNKNSFLVRESQKPGDFVLVSRDQKVTVMIRWQKIVDGGESF
GTLSELIDYKKNPMVETCTVHLRQPNFNAITAAGINARVEQLVKGVEFEFESL
QDSRTFRNKGYSKQENLKNRYNLPDTRVRLKLDVEHSAVAGAEINANTIRL
TGDDLXMSSESLSNSVSPCPACTAQTQRCNSCLOQKTCVQCAVKAILPYSPN
CATCRKSDLSLKHKSESASSPSGSGSGSGTSGSGTGGTGGTGGTGGTGGTGGT
GCLVGLKRNDSGAVISMAEREREREMFXYATQGCLLTQVNTDPMWQENTRV
VMOENTRVIMTKYERKEKCAVWPDGHPADPCVLNPLQVNTROSHLAQERKPGI
VSWRDQPARFIHFQVWPDGHPADPCVLNPLQVNTROSHLAQERKPGI
CSAGIGTGTGTIVIMDLQIVRNGDITDITQRTQVRSQSLQVREYTNKYTGIE
AVUHYQTLIAKRAEQSLQVREYTNKYTGIEGNSORSPLPAPTSSILVSPKLT
PTPTSDTLGTSGLSMGVGMGNKHSKQPPPLPVNCCNNGNIGNSCSGSGGS
STTSSNGSNGINALGGIGLGGNNRKNFYSLSLQKQOREQAPAGAKWQFAP
PAPLPRPGILKLTSPVIFQNSKTFPKT"
PAPLPRPGILKLTSPVIFQNSKTFPKT"
ORIGIN
Query Match 3.6%; Score 39.6; DB 3; Length 5513;
Best Local Similarity 52.4%; Pred. No. 15;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 569 TTGAGGCTATCTCAGACCGAGGAAAGTTGGCGCACTGGCAAGCGTAGTTGCCACGG 628
Db 3979 TTAAGACCTACATCGCCACCCAGGGGTGTCTGTCTACCCAGCAAGTGAACCGGTACGG 4038
Qy 629 AGGCACGACGACGAGCGCGGCGGACCAACCCCGCGCGCTCACTCAACACCGCGCCT 688
Db 4039 ACTTTGGAACATGGTCTGGCAGAGAACACGCGGGTGATCGTCATGACACCAAGAGT 4098
Qy 689 CCGAAGCTTGAAGCGGACACCTTTACCTTCCCTCCCTGGGTAGAAAATGG 734
Db 4099 ACGAGCGCGCGCAAGAAAGTGGCGCGCTACTTGGCCAGCAGAGGG 4144
```

```
gene
/notes="genotype: isofemale"
<161..>.5450
/genes="csw"
join(<161..174,757..944,1017..1143,2097..2259,2952..3131,
3200..3317,3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/product="corkscrew phosphatase splice variant A"
join(161..174,757..944,1017..1143,2097..2259,2952..3131,
3200..3317,3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/codon_start=1
/product="corkscrew phosphatase splice variant A"
/protein_id="AAN17619.1"
/db_xref="GI:23344954"
/translation="MSRRWFHTISGIEAEKLBQPGDGSFLARLSSNPCATLIS
VRNGEVTHIKIQQNDFDLVYGGKEFATLPELVYIMENGELKKGQAIELKQPLI
CAEPTTERFHGSLGKEAEKLIILERKNGSLFVRESQSGPGFVLSVTRDDKTVTHM
IRWQKKYDVGGESFTLSLIDHYKRNPMVETCGTVVHLRQPFNATRIITAAGINAR
VEOLVKGFWEEFESLQDSRDTFSENEGYKOENLKNEYRNILPYDHTRVKLLDVEH
SVAGAYINAYIRLFTDGLNYMSSSESLSNVSPSCPACTAATQRCNCQQLQNK
TCVQAVKAILPYSNCACTSRKSDLSLKHKSSESASSPSSGSGSGSGSGTSGVS
SVNPGPTNLTGSLAGCLVGLLKRHSNDSGAVSISMAEREREREMFKTYIATQG
CLLTQOQNTVTDFNNWMOENTRVIMTKEYERGKECARYWDEGRSEOFGHARIO
CVSENSTDYTLREELVSWDOPARRIHYHFQWPDHGVADPCVLNLFLODVNTRQ
SHLAQAGEPGPICVCSAGIGRTGTFVIDMILDOIVRGLDTEIDIQRTIQMVRSQ
RSLVQTEAQYKEVYAVQHYIOTLIARKRAEQSLQVGREYTNIKYTGIEIGNSQRS
PLPAISSTSLPVSKEPTLPTADLTGSLGMVGMGNKSHASKQPPLEPVNCCN
NNGIIGNSCSNGSGSTSSSSNGNINALLGGIIGLGNMKNKSNFYSDSLKQQ
QOREQAPAGAKMQPAPPLRPRGILKLTSPVIFQNSKTPPKT"
join(<1332..1960,2097..2259,2952..3131,3200..3317,
3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
join(1332..1960,2097..2259,2952..3131,3200..3317,
3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
/protein_id="AAN17620.1"
/db_xref="GI:23344955"
/translation="MLFNKLEKLSLGLNVNHNKLOEKQVYNNNNNNNNNNNNNN
NATYNNQNEYEAIQAHYSGKRRSEERERSGKFKASGRKAKVTPPTPEAOEPA
CKNCMTHELAQIIGVAGKADQRNRRLRPLSAQPSAASASTESLHRL
TPSQASYATPTPTPPQAPFGASCNSLTLATMRVQLHGYTWPHGNLSG
KEAEKILIRGKNGSLVRESQSGPGFVLSVTRDDKTVHMRWQDKYDVGGESF
GTLSELIDHYKRNPMVETCGTVVHLRQPFNATRIITAAGINAEVQLVKGFWEEFESL
QQSRDTPFRNEGYKQENLKNEYRNILPYDHTRVKLLDVHSAEYINAYIRLP
TDGDLNMSSESLSNVSPSCPACTAATQRCNCQQLQNKTCVCAVKSAILPYSN
CATCSRKSDLSLKHKSSESASSPSSGSGSGSGTSGVSVNGPPTNLTGSGTA
GCLVGLLKRHSNDSGAVSISMAEREREREMFKTYIATQGLLTQOQNTVTDFNNW
VMQENTRVIMTKEYERGKECARYWDEGRSEOFGHARIOCVSENSTDYTLREEL
VSWRDPARRIHYHFQWPDHGVADPCVLNLFLODVNTRQSHLAQAGEKPGPICVH
CSAGIGRTGTFVIDMILDOIVRGLDTEIDIQRTIQMVRSORSLVQTEAQYKEVY
AVQHYIOTLIARKRAEQSLQVGREYTNIKYTGIEIGNSQRSPLPAISSTSLVPSKT
PLTPTADLTGSLGMVGMGNKSHASKQPPLEPVNCCNNGIIGNSCSNGSGS
STTSSNGSGNINALLGGIIGLGNMKNKSNFYSDSLKQQQREEQAPAGAKRWQQ
PAPPLRPRGILKLTSPVIFQNSKTPPKT"
```

ORIGIN

```
Query Match          3.6%; Score 39.6; DB 3; Length 5514;
Best Local Similarity 52.4%; Pred. No. 15;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 569 TTGAGGCCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3980 TTAAGACCTACATCGCCACCCAGCGGTGCTGCTGCCACCAAGTGAACACGGTGACGG 4039
QY 629 AGGCCAGCAGCAGCGCGGAGCGAGCAACCCGACGCCCGCTACTGCAACAGCGCCT 688
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4040 ACTTCTGGAACATGCTGTCAGGAGAACACGCGGGGTGATCGTCATGACCACCAAGAGT 4099
QY 689 CCGAACTTGAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 20:07:34 ; Search time 552 Seconds  
(without alignments)  
8357.868 Million cell updates/sec

Title: US-09-993-777-1\_COPY\_9874\_10959  
Perfect score: 1086  
Sequence: 1 ATGCATCGTACTTGGAC.....CGCCGACTCATAAAATG 1086

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Ab106707 Drosophil  
Ab106706 Drosophil  
Aax84332 Stealth v  
Aax10594 Gene enco  
Aat44849 Mycelioph  
Abq37730 Oligonuc  
Abq37731 Oligonuc  
Abx07675 S. pneumo  
Aah30774 CPE 78 co  
Abz42309 Streptoco  
Aah30883 2CFE 78 c  
Aaz50960 A. haloph  
Aav52232 Streptoco  
Continuation (17 o  
Aas93325 DNA encod  
Aas8627 DNA encod  
Aas90013 DNA encod  
Aav88421 EST clone  
Ach34064 Human end  
Ab199754 Human sec  
ACC90579 Human CGD  
Abx76252 Lung canc

ALIGNMENTS

RESULT 1  
AAT33504  
ID AAT33504 standard; DNA; 13473 BP.  
XX AAT33504;  
AC AAT33504;  
XX XX  
DT 16-OCT-2003 (revised)  
DT 18-NOV-1996 (first entry)  
XX DE Infectious laryngotracheitis virus short region sequence.  
XX XX  
XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;  
XX vector; vaccine; chicken; poultry; immunisation; ds.  
XX Gallid herpesvirus 1.  
XX XX  
FH Key Location/Qualifiers  
FT repeat\_region 1..273  
FT /tag= a "internal repeat region"  
FT misc\_feature 274..13371  
FT /function= "unique short sequence of ILTV"  
FT CDS complement(281..970)  
FT /tag= c  
FT /label= US2\_gene  
FT CDS 1059..2489  
FT /tag= d  
FT /label= Protein-kinase\_gene  
FT CDS 2575..4107  
FT /tag= e  
FT /label= UL47-like\_gene  
FT CDS 4113..4445  
FT /tag= f  
FT /label= ORF4  
FT CDS complement(4139..4519)  
FT /tag= g  
FT /label= ORF4\_reverse\_complement  
FT misc\_difference 4535  
FT /tag= h  
FT /note= "base 4535 is given as 's' in the specification"  
FT CDS 4609..5487  
FT /tag= i  
FT /label= gg\_gene  
FT CDS 5697..8654  
FT /tag= j

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	13473	2	AAT33504 Infectiou
2	1086	100.0	13473	2	AAT44384 Infectiou
3	1084.4	99.9	3605	6	ABK90557 Infectiou
4	1084.4	99.9	18912	2	AAT44385 Infectiou
C 5	343	31.6	534	2	AAT33513 Infectiou
C 6	47.8	4.4	2000	7	ADA71938 Rice gene
7	41.4	3.8	2000	7	ADA71938 Rice gene
C 8	38.2	3.5	777	7	ACA38608 Prokaryot
C 9	38.2	3.5	780	7	ACA40341 Prokaryot
C 10	38.2	3.5	110000	4	AAI199682_06 Continuation (7 of
C 11	38.2	3.5	110000	4	AAI199683_06 Continuation (7 of
12	38	3.5	2526	5	AAT26455 D.melanog
13	38	3.5	2677	4	AB106927 Drosophil
14	38	3.5	2838	4	AB118521 Drosophil
15	38	3.5	4528	4	AB121135 Drosophil
16	38	3.5	10587	4	AB118520 Drosophil
17	38	3.5	19468	4	AB106926 Drosophil
18	38	3.5	21399	4	AB121134 Drosophil
19	37.6	3.5	813	6	ABQ13969 Oligonuc
C 20	37.6	3.5	813	6	ABQ13968 Oligonuc
C 21	37.6	3.5	6310	6	ABK31365 Signal tr
C 22	37.6	3.5	6310	6	AB170322 Chemical
C 23	37.6	3.5	6310	6	AAS61269 Human gen

FT CDS /label= g60\_gene  
FT complement(6948..7826)  
FT /\*tag= k  
FT /label= ORF6\_reverse\_complement  
FT 8462..9766  
FT /\*tag= 1  
FT /label= gD\_gene  
FT 9874..10962  
FT /\*tag= m  
FT /label= gI\_gene  
FT complement(10617..11150)  
FT /\*tag= n  
FT /label= ORF8\_reverse\_complement  
FT 11159..12658  
FT /\*tag= o  
FT /label= gE\_gene  
FT 12665..13447  
FT /\*tag= p  
FT /label= ORF10  
FT misc\_difference 13002..13003  
FT /\*tag= q  
FT /notes "bases 13002-13303 are given as 'ss' in the  
FT specification"  
FT repeat\_region 13372..13473  
FT /\*tag= r  
FT /function= "terminal repeat region"  
FT  
FT W09508622-A1.  
FT  
FT 30-MAR-1995.  
FT  
FT 16-SEP-1994; 94WO-US010628.  
FT  
FT 24-SEP-1993; 93US-00126597.  
FT  
FT (SYTR ) SYNTR0 CORP.  
FT  
FT Wild MA, Cochran MD;  
FT  
FT WPI; 1995-139591/18.  
FT  
FT P-PSDB; AAW00630, AAW00631, AAW00632, AAW00633, AAW00634, AAW00635,  
FT AAW00636, AAW00637, AAW00638, AAW00639, AAW00640, AAW00641, AAW00642.  
FT  
FT Recombinant attenuated infectious laryngotracheitis virus - for use in  
FT vaccines to protect poultry from infection from the virus, also methods  
FT of distinguishing between vaccinated and naturally infected birds.  
FT  
FT Example 1; Page 79-94; 177pp; English.  
FT  
FT The unique short region (AAT33504) of infectious laryngotracheitis virus  
FT (ILT) genomic DNA contains genes (see also AAT33505 and AAT33510-13)  
FT that are associated with ILTV virulence. A deletion in those genes, esp.  
FT the glycoprotein gG gene, glycoprotein gI gene, thymidine kinase gene,  
FT US2 gene, UL47-like gene or the glycoprotein g60 gene, will attenuate the  
FT ILTV. A gene for a foreign antigen may be inserted into the US2, UL47-  
FT like, ORF4, gG, g60 or gI gene to produce a recombinant ILTV. Recombinant  
FT ILTV may be used as a multivalent vaccine, esp. for use in poultry.  
FT (Updated on 16-OCT-2003 to standardise OS field)  
FT  
FT Sequence 13473 BP; 3390 A; 3582 C; 3547 G; 2951 T; 0 U; 3 Other;  
FT  
FT Query Match 100.0%; Score 1086; DB 2; Length 13473;  
FT Best Local Similarity 100.0%; Pred. No. 0;  
FT Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
FT  
FT 1 ATGGATCGTACTGGAATCTGGCTCTCCCTGGCGGACGCTGACCCCTTCGGCGG 60  
FT 9874 ATGGATCGTACTGGAATCTGGCTCTCCCTGGCGGACGCTGACCCCTTCGGCGG 9933  
FT  
FT 61 ATGGGAATCGTATCACTGGAATCACTGCTCCGCCAGGATTGACGAGTACATCGTG 120  
FT 9934 ATGGGAATCGTATCACTGGAATCACTGCTCCGCCAGGATTGACGAGTACATCGTG 9993  
FT

Qy 121 ATGTCGCGCTCGCCCGAAGCTCAATTAACCTGAGCTATTTTTCATGCTGCTGCGCAG 180  
Db 9994 ATGTCGCGCTCGCCCGAAGCTCAATTAACCTGAGCTATTTTTCATGCTGCTGCGCAG 10053  
Qy 181 AGACCCCAAAACCTTACTCAGGAACCGTCGCGCTTCGCTGCTGATATAAACAAC 240  
Db 10054 AGACCCCAAAACCTTACTCAGGAACCGTCGCGCTTCGCTGCTGATATAAACAAC 10113  
Qy 241 CAGTGCTACCAAGAACTTAGCGAGGAGCGCTTTGAAAATTTGCACTCATCGATCGTCTTCT 300  
Db 10114 CAGTGCTACCAAGAACTTAGCGAGGAGCGCTTTGAAAATTTGCACTCATCGATCGTCTTCT 10173  
Qy 301 GTTTTGTGCGCTGTAAGTGACCGAGTACACGTTCTCCGCTCGAAGCACTAACCGGA 360  
Db 10174 GTTTTGTGCGCTGTAAGTGACCGAGTACACGTTCTCCGCTCGAAGCACTAACCGGA 10233  
Qy 361 CTTCCACACCGCTTTAAGCTCACTATACGAATCTCTGCTCCGAACGACACGCGGATGTTT 420  
Db 10234 CTTCCACACCGCTTTAAGCTCACTATACGAATCTCTGCTCCGAACGACACGCGGATGTTT 10293  
Qy 421 TAGCTAATTTGCTGGCTAGACGACCAAAAGAACCCATTGACGCTTCGCGATCCAACTA 480  
Db 10294 TAGCTAATTTGCTGGCTAGACGACCAAAAGAACCCATTGACGCTTCGCGATCCAACTA 10353  
Qy 481 TCGGTGTATCAATTCGGGAACACCGCGGAGCTCGGGACTCTATTCCAGGCTTCGTGT 540  
Db 10354 TCGGTGTATCAATTCGGGAACACCGCGGAGCTCGGGACTCTATTCCAGGCTTCGTGT 10413  
Qy 541 CGCACCTTCGGATTACCTACCGTCCAACTTTAGCGCTTATCTCAGGACCGAGGAAGTTGG 600  
Db 10414 CGCACCTTCGGATTACCTACCGTCCAACTTTAGCGCTTATCTCAGGACCGAGGAAGTTGG 10473  
Qy 601 CGCAACTGGAAGCTGACGTTGCCACGAGGCCACGACGACGCGCGGAGGCGACAACC 660  
Db 10474 CGCAACTGGAAGCTGACGTTGCCACGAGGCCACGACGACGCGCGGAGGCGACAACC 10533  
Qy 661 CGGACGCGCTCTGCAACACGCGCTTCGAACTTTGAAGCGGAACACTTTTACCTTTCCC 720  
Db 10534 CGGACGCGCTCTGCAACACGCGCTTCGAACTTTGAAGCGGAACACTTTTACCTTTCCC 10593  
Qy 721 TGGCTAGAAATGCGCTGGATCATTAACGACGACACCGCGCAACGAAATTCAAACGTT 780  
Db 10594 TGGCTAGAAATGCGCTGGATCATTAACGACGACACCGCGCAACGAAATTCAAACGTT 10653  
Qy 781 ACTGTCGCTCGCGCAATGAGCCCTTACGCTAATTTGGGGTAACCGTGGCTGCGCTCGTG 840  
Db 10654 ACTGTCGCTCGCGCAATGAGCCCTTACGCTAATTTGGGGTAACCGTGGCTGCGCTCGTG 10713  
Qy 841 AGCGCAACGATCGCGCTCGCTCATTTGTAATTTCCATCGTCAACGAAACATGTGACCCCG 900  
Db 10714 AGCGCAACGATCGCGCTCGCTCATTTGTAATTTCCATCGTCAACGAAACATGTGACCCCG 10773  
Qy 901 CACGAAAATTAGACACGCTTCGAGAGCAGCAGAGAGAGCTTCCCAACTAGAGGGAA 960  
Db 10774 CACGAAAATTAGACACGCTTCGAGAGCAGCAGAGAGAGCTTCCCAACTAGAGGGAA 10833  
Qy 961 TCGGAAAATTTGACCCATGTTGGTGGGTAATTAACGAGGCGCTGACCAAGGATG 1020  
Db 10834 TCGGAAAATTTGACCCATGTTGGTGGGTAATTAACGAGGCGCTGACCAAGGATG 10893  
Qy 1021 GAATCTGGAACCTGGTTGGATTGTTAAACCGCTCTGCGCTGAGCTGCGCCGACTCAATA 1080  
Db 10894 GAATCTGGAACCTGGTTGGATTGTTAAACCGCTCTGCGCTGAGCTGCGCCGACTCAATA 10953  
Qy 1081 AAAATG 1086  
Db 10954 AAAATG 10959

RESULT 2  
AAT44384  
ID AAT44384 standard; DNA; 13473 BP.  
XX



PT Recombinant infectious laryngotracheitis virus with deletion in the  
PT glyco:protein G, g1 or US2 gene, etc. - useful for vaccines against  
PT infectious laryngotracheitis in poultry.  
XX Example 11; Page 88-103; 216pp; English.  
PS  
XX The nucleotide sequence of 13,473 bp of contiguous DNA (AAT44384) from  
CC the unique short region of infectious laryngotracheitis virus (ILTV)  
CC contains the entire 13,098 bp unique short region and includes 13 open  
CC reading frames that encode proteins (AAM06782-94) of over 100 amino  
CC acids, 8 of which show significant homology to other virus genes. Novel  
CC recombinant, attenuated ILTV comprises the ILTV genome contg. a deletion  
CC in the unique short region, esp. in the glycoprotein gg, g1, US2, ORF4,  
CC UL47-like or g60 gene. The attenuated virus is useful as a vaccine  
CC against ILTV. A foreign gene encoding an antigen e.g. from another avian  
CC virus can be inserted into the US2, UL47-like, ORF4, g60 or g1 gene  
CC to provide a multivalent vaccine for chickens and other poultry. Deletion  
CC of the gg or g1 gene provides as a negative marker to distinguish  
CC vaccinated from infected animals. (Updated on 16-OCT-2003 to standardise  
CC OS field)  
XX  
SQ Sequence 13473 BP; 3390 A; 3581 C; 3547 G; 2952 T; 0 U; 3 Other;  
Query Match 100.0%; Score 1086; DB 2; Length 13473;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGATCGCTACTTGGAACTCTGGCTCTCTTGGCGGAGCGCTCGACCCCTTGGCGCG 60  
DB 9874 ATGGGATCGCTACTTGGAACTCTGGCTCTCTTGGCGGAGCGCTCGACCCCTTGGCGCG 9933  
QY 61 ATGGGATCGCTACTTGGAACTCTGGCTCTCTTGGCGGAGCGCTCGACCGATCAGATCGTG 120  
DB 9934 ATGGGATCGCTACTTGGAACTCTGGCTCTCTTGGCGGAGCGCTCGACCGATCAGATCGTG 9993  
QY 121 ATCGTCGCGCTCTGCGCCGAGCTCAATTCAGTGCAGTATTTTTCATGCTGGCCAG 180  
DB 9994 ATCGTCGCGCTCTGCGCCGAGCTCAATTCAGTGCAGTATTTTTCATGCTGGCCAG 10053  
QY 181 AGACCCCAACAAACCTTACTCAGAAACCGTCCGCGTGGCTTTTCGCTCTGATATACAAAC 240  
DB 10054 AGACCCCAACAAACCTTACTCAGAAACCGTCCGCGTGGCTTTTCGCTCTGATATACAAAC 10113  
QY 241 CAGTGTACAGAACTTACGAGGAGCGCTTTGAAATGACATCAGTCTGCTCTCT 300  
DB 10114 CAGTGTACAGAACTTACGAGGAGCGCTTTGAAATGACATCAGTCTGCTCTCT 10173  
QY 301 GTTTTGTGCGTGTAAAGTACCGAGTACAGTCTTCGCGCTCGAAGACAGTCAACCGGA 360  
DB 10174 GTTTTGTGCGTGTAAAGTACCGAGTACAGTCTTCGCGCTCGAAGACAGTCAACCGGA 10233  
QY 361 CTCCACACCCGTTTAAAGTCTACTATACGAAATCTCTCGTCCGAAACGACGCGGATGTC 420  
DB 10234 CTCCACACCCGTTTAAAGTCTACTATACGAAATCTCTCGTCCGAAACGACGCGGATGTC 10293  
QY 421 TAGCTAATTTGTCGGTGTAGACGACACCAAGAACCCATTCAGTCTTCGCGATCCAACTA 480  
DB 10294 TAGCTAATTTGTCGGTGTAGACGACACCAAGAACCCATTCAGTCTTCGCGATCCAACTA 10353  
QY 481 TCGGTGTATCAATTCGGAACACCGCGCGGACCTCGCGGAGCTCTATTCGAGGCTTCGTGT 540  
DB 10354 TCGGTGTATCAATTCGGAACACCGCGCGGAGCTCTATTCGAGGCTTCGTGT 10413  
QY 541 CGCACCTTCGGATTACTACCGTCCAACTTGGAGGCTCTATTCAGGACCGAGGAAAGTTGG 600  
DB 10414 CGCACCTTCGGATTACTACCGTCCAACTTGGAGGCTCTATTCAGGACCGAGGAAAGTTGG 10473  
QY 601 CGCAACTGGCAAGCGTACGTTGCGACGAGGCGCACGACGACGACGCGCGGAGCGCAAC 660  
DB 10474 CGCAACTGGCAAGCGTACGTTGCGACGAGGCGCACGACGACGACGCGCGGAGCGCAAC 10533  
QY 661 CCGACGCGCGTCACTGCAACACGCGCTTCGAACTTGAAGCGGAAACACTTACCTTCC 720

DB 10534 CCGACGCCGCTCACTGCAACACCGCCTCCGAACTTGAAGCGGAAACACTTTACCTTCC 10593  
QY 721 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACACCGCGAAACGAAAATTCAAACGTT 780  
DB 10594 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACACCGCGAAACGAAAATTCAAACGTT 10653  
QY 781 ACTGTCGCTCTCGGAGCAATGAGCCCTACGCTAATTTGGGTAACCGTGGCTGCCGTCGTG 840  
DB 10654 ACTGTCGCTCTCGGAGCAATGAGCCCTACGCTAATTTGGGTAACCGTGGCTGCCGTCGTG 10713  
QY 841 AGCGCAACGATCGGCTCTCGCTCAATTTGTAATTTCCATCGTCCACGAAACATGTGCAACCG 900  
DB 10714 AGCGCAACGATCGGCTCTCGCTCAATTTGTAATTTCCATCGTCCACGAAACATGTGCAACCG 10773  
QY 901 CACCGAAAATTAGACACCGTCTCGCAAGACGACGAAACGTTCCCAAACTAGAGGAGAA 960  
DB 10774 CACCGAAAATTAGACACCGTCTCGCAAGACGACGAAACGTTCCCAAACTAGAGGAGAA 10833  
QY 961 TCGGAAAATTTGGACCCATGTTGCTGCGGAAATAAACAAGGCGCTGACCAAGATAGT 1020  
DB 10834 TCGGAAAATTTGGACCCATGTTGCTGCGGAAATAAACAAGGCGCTGACCAAGATAGT 10893  
QY 1021 GAATTTGTGGAAGTGGTTCGGATTTGTAACCGCTCTGCGCTAAGCTTCGCCCGACTCAATA 1080  
DB 10894 GAATTTGTGGAAGTGGTTCGGATTTGTAACCGCTCTGCGCTAAGCTTCGCCCGACTCAATA 10953  
QY 1081 AAAATG 1086  
DB 10954 AAAATG 10959  
RESULT 3  
ABK90557  
ID ABK90557 standard; DNA; 3605 BP.  
XX AC ABK90557;  
XX DT 29-AUG-2003 (revised)  
XX DT 15-NOV-2002 (first entry)  
XX Infectious laryngotracheitis virus (ILTV) DNA.  
DE Infectious laryngotracheitis virus; novel recombinant avian herpesvirus;  
KW NAHV; herpes virus of turkeys; HVT; Marek's disease; Newcastle disease;  
KW infectious laryngotracheitis; virucide; immunostimulant; vaccine; ILTV;  
KW gene; ds.  
XX Gallid herpesvirus 1.  
OS  
XX Location/Qualifiers  
FH 585..1889  
FT /\*tag= a  
FT /product= "ILTV glycoprotein D"  
FT 1997..3085  
FT /\*tag= b  
FT /product= "ILTV glycoprotein I"  
XX US2002081316-A1.  
XX  
XX 27-JUN-2002.  
XX  
XX 14-JUN-2001; 2001US-00881457.  
XX  
XX 12-JUN-1992; 92US-00898087.  
PR 26-FEB-1993; 93US-00023610.  
PR 14-JUN-1993; 93WO-US005681.  
PR 09-AUG-1994; 94US-00288065.  
PR 09-AUG-1995; 95WO-US010245.  
PR 13-JUN-1996; 96US-00663566.  
PR 21-FEB-1997; 97US-00804372.  
PR 25-OCT-1999; 99US-00426352.  
XX (COCH/) COCHRAN M D.  
PA





```

FT FT      /tag= f
FT FT      /label= ORF1(RC)
FT FT      /note= "unique-like short 2 (US2) gene"
FT FT      3642..3645
FT FT      /tag= g
FT FT      /note= "TATA signal for ORF1 and ORF2"
FT FT      3675..3678
FT FT      /tag= h
FT FT      /note= "TATA signal for ORF1"
FT FT      3677..3680
FT FT      /tag= i
FT FT      /note= "TATA signal for ORF2"
FT FT      3694..5124
FT FT      /tag= j
FT FT      /label= ORF2
FT FT      /note= "protein kinase gene"
FT FT      5125..5131
FT FT      /tag= k
FT FT      /note= "polyA signal for ORF2"
FT FT      5173..5178
FT FT      /tag= l
FT FT      /note= "polyA signal for ORF2"
FT FT      5210..7081
FT FT      /tag= m
FT FT      /label= ORF3
FT FT      /note= "unique long 47 (UL47)-like gene"
FT FT      6748..7080
FT FT      /tag= o
FT FT      /label= ORF4
FT FT      /note= "complement(6774..7154)"
FT FT      /tag= p
FT FT      /label= ORF4(RC)
FT FT      7158..7161
FT FT      /tag= n
FT FT      /note= "TATA signal for ORF4"
FT FT      7158..7161
FT FT      /tag= q
FT FT      /note= "TATA signal for ORF5"
FT FT      7245..8123
FT FT      /tag= r
FT FT      /label= ORF5
FT FT      /note= "glycoprotein gG gene"
FT FT      7245..7322
FT FT      /tag= s
FT FT      7323..8120
FT FT      /tag= t
FT FT      8200..8205
FT FT      /tag= u
FT FT      /note= "polyA site for ORF5"
FT FT      8289..8294
FT FT      /tag= v
FT FT      /note= "polyA signal for ORF5"
FT FT      8333..11290
FT FT      /tag= w
FT FT      /label= ORF6
FT FT      /note= "glycoprotein g60 gene"
FT FT      complement(9584..10462)
FT FT      /tag= y
FT FT      /label= ORF6(RC)
FT FT      9623..10363
FT FT      /tag= x
FT FT      /note= "repeat region consists of approx. 23 repeats of
FT FT      30-36 bp"
FT FT      11091..11094
FT FT      /tag= z
FT FT      /note= "potential TATA signal for ORF7"
FT FT      11096..12402
FT FT      /tag= aa
FT FT      /label= ORF7
FT FT      12455..12448
FT FT      /tag= ab
FT FT      /note= "TATA signal for ORF8"
FT FT      12510..13598

```

```

FT FT      /tag= ac
FT FT      /label= ORF8
FT FT      /note= "glycoprotein gI gene"
FT FT      12510..12575
FT FT      /tag= ad
FT FT      12576..13595
FT FT      /tag= ae
FT FT      complement(13253..13786)
FT FT      /tag= af
FT FT      /label= ORF8(RC)
FT FT      13705..13708
FT FT      /tag= ag
FT FT      /note= "TATA signal for ORF9"
FT FT      13792..15291
FT FT      /tag= ah
FT FT      /label= ORF9
FT FT      /note= "glycoprotein gE gene"
FT FT      13792..13845
FT FT      /tag= ai
FT FT      13846..15288
FT FT      /tag= aj
FT FT      15116..15119
FT FT      /tag= ak
FT FT      /note= "TATA signal for ORF10"
FT FT      15298..16080
FT FT      /tag= al
FT FT      /label= ORF10
FT FT      16004..18912
FT FT      /tag= am
FT FT      16129..17013
FT FT      /tag= an
FT FT      complement(17380..18216)
FT FT      /tag= ao
FT FT      WO9629396-A1.
FT FT      XX
FT FT      PD 26-SEP-1996.
FT FT      XX
FT FT      PF 21-MAR-1996; 96WO-US003916.
FT FT      XX
FT FT      PR 23-MAR-1995; 95US-00410121.
FT FT      PR 06-JUN-1995; 95US-00468190.
FT FT      XX
FT FT      PA (SYTR ) SYNTRO CORP.
FT FT      XX
FT FT      PI Wild MA, Cochran MD;
FT FT      XX
FT FT      DR WPI; 1996-443172/44.
FT FT      DR P-PSDB; AAW06782, AAW06783, AAW06784, AAW06785, AAW06786, AAW06787,
FT FT      DR AAW06788, AAW06789, AAW06790, AAW06791, AAW06792, AAW06793, AAW01415,
FT FT      DR AAW01416.
FT FT      XX
FT FT      PT Recombinant infectious laryngotracheitis virus with deletion in the
FT FT      glyco:protein G, gI or US2 gene, etc. - useful for vaccines against
FT FT      infectious laryngotracheitis in poultry.
FT FT      XX
FT FT      PS Example 11; Page 138-154; 216pp; English.
FT FT      XX
FT FT      CC The nucleotide sequence of 19,912 bp of contiguous DNA (AAU44385) from
FT FT      the unique short and flanking region of infectious laryngotracheitis
FT FT      virus (ILT) contains the entire 13,098 bp unique short region (see also
FT FT      AAU44384) and includes 17 open reading frames that encode proteins
FT FT      (AAW06782-94, AAW01415-16) of over 100 amino acids, 10 of which show
FT FT      homology to other virus genes. Novel recombinant, attenuated ILTV
FT FT      comprises the ILTV genome contg. a deletion in the unique short region,
FT FT      esp. in the glycoprotein gG, gI, US2, ORF4, UL47-like or g60 gene. The
FT FT      attenuated virus is useful as a vaccine against ILTV. A foreign gene
FT FT      encoding an antigen e.g. from another avian virus can be inserted into
FT FT      the US2, UL47-like, ORF4, gG, g60 or gI gene to provide a multivalent
FT FT      vaccine for chickens and other poultry. Deletion of the gG or gI gene
FT FT      provides as a negative marker to distinguish vaccinated from infected
FT FT      animals. (Updated on 16-OCT-2003 to standardise OS field)
FT FT      XX

```

SQ Sequence 18912 BP; 4416 A; 5261 C; 5251 G; 3984 T; 0 U; 0 Other;  
Query Match 99.9%; Score 1084.4; DB 2; Length 18912;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCATCGCTACTTGGAACTCTGGCTCTCTTCCGCGAGCTCGACCTTCGGGCG 60  
DB 12510 ATGGCATCGCTACTTGGAACTCTGGCTCTCTTCCGCGAGCTCGACCTTCGGGCG 12569  
QY 61 ATGGGAATCGTGATCACTGGAAATCACGCTTCCGCGAGGATGACGAGATCAATCGTG 120  
DB 12570 ATGGGAATCGTGATCACTGGAAATCACGCTTCCGCGAGGATGACGAGATCAATCGTG 12629  
QY 121 ATCGTCGGCGCTCGCCCGAGCTACAATCAACTGCAGCTATTTTCATCGCTGGCCAG 180  
DB 12630 ATCGTCGGCGCTCGCCCGAGCTACAATCAACTGCAGCTATTTTCATCGCTGGCCAG 12689  
QY 181 AGACCCCAAAACCTTACTCAGAAACCTCGCGCTTCGCTCTGATATACAAAC 240  
DB 12690 AGACCCCAAAACCTTACTCAGAAACCTCGCGCTTCGCTCTGATATACAAAC 12749  
QY 241 CAGTCTACCAAGAACTTAGCGAGGCGCTTTGAAAATTCGACTCATCGATCGTCTCT 300  
DB 12750 CAGTCTACCAAGAACTTAGCGAGGCGCTTTGAAAATTCGACTCATCGATCGTCTCT 12809  
QY 301 GTTTTGTGCGTGAAGTACCGAGTACAGTTCTCGCGCTCGAAACAGACTAACCGGA 360  
DB 12810 GTTTTGTGCGTGAAGTACCGAGTACAGTTCTCGCGCTCGAAACAGACTAACCGGA 12869  
QY 361 CTTCCACACCCGTTTAAGTCTACTATAGAAATCTCTCGCGAGCAGCGGATGTTTC 420  
DB 12870 CTTCCACACCCGTTTAAGTCTACTATAGAAATCTCTCGCGAGCAGCGGATGTTTC 12929  
QY 421 TACGTAATTTGTCGCTAGACACACCAAGAACCCATTGACGCTTCGCGATCCAACTA 480  
DB 12930 TACGTAATTTGTCGCTAGACACACCAAGAACCCATTGACGCTTCGCGATCCAACTA 12989  
QY 481 TCGGTGTATCAATTCGCGAACCCTCGCGACTTCGCGACTTATTCAGAGCTTCGTTGT 540  
DB 12990 TCGGTGTATCAATTCGCGAACCCTCGCGACTTCGCGACTTATTCAGAGCTTCGTTGT 13049  
QY 541 CGACCTTCGATACCTACCTGCTCAACTGAGGCTATCTCAGACCGAGGAAGTTGG 600  
DB 13050 CGACCTTCGATACCTACCTGCTCAACTGAGGCTATCTCAGACCGAGGAAGTTGG 13109  
QY 601 CGCACTGCAAGCTGCTTCCGCGAGGCGCAGCAGCAGCGCGCGAGCAGCAGC 660  
DB 13110 CGCACTGCAAGCTGCTTCCGCGAGGCGCAGCAGCAGCGCGCGAGCAGCAGC 13169  
QY 661 CGAGCGCGCTCACTGCAACGAGCGCTCGGAACTTGAAGCGGAACATTTACCTTTCCC 720  
DB 13170 CGAGCGCGCTCACTGCAACGAGCGCTCGGAACTTGAAGCGGAACATTTACCTTTCCC 13229  
QY 721 TGGCTAGAAAATGGCGTGATCATTCAGAACCGACCGCGCAACCGAAATTCAAAGTT 780  
DB 13230 TGGCTAGAAAATGGCGTGATCATTCAGAACCGACCGCGCAACCGAAATTCAAAGTT 13289  
QY 781 ATGTCCGCTCGCGAACAATGAGCCCTAGCTTAATTTGGGTHAACCGTGGCTGCGTCTGT 840  
DB 13290 ATGTCCGCTCGCGAACAATGAGCCCTAGCTTAATTTGGGTHAACCGTGGCTGCGTCTGT 13349  
QY 841 AGCGCAACGATGGCGCTGCTGATTTGTAATTTCCATCTGTCACAGAAACATGTGACCCCG 900  
DB 13350 AGCGCAACGATGGCGCTGCTGATTTGTAATTTCCATCTGTCACAGAAACATGTGACCCCG 13409  
QY 901 CACCGAAAATTAGACAGGCTCTCGCAAGACGACGAGAACGTTTCCCAACTAGAGGGAA 960  
DB 13410 CACCGAAAATTAGACAGGCTCTCGCAAGACGACGAGAACGTTTCCCAACTAGAGGGAA 13469  
QY 961 TCGCGAAAATTTGGACCCCATGTTTGGTCTCGAAAATAAACAAGGGGCGCTGACCAAGATAGT 1020  
DB 13470 TCGCGAAAATTTGGACCCCATGTTTGGTCTCGAAAATAAACAAGGGGCGCTGACCAAGATAGT 13529

QY 1021 GAACTTGTGAACTGGTTCGGAATTGTAACCGCTTCGCGTAAGCTCGCCGACTCAATA 1080  
DB 13530 GAACTTGTGAACTGGTTCGGAATTGTAACCGCTTCGCGTAAGCTCGCCGACTCAATA 13589  
QY 1081 AATAATG 1086  
DB 13590 AATAATG 13595

## RESULT 5

AAT33513/c  
ID AAT33513 standard; DNA; 534 BP.

XX AAT33513;  
XX AC  
XX 16-OCT-2003 (revised)  
DT 19-NOV-1996 (first entry)  
XX Infectious laryngotracheitis virus ORF8 reverse complement.  
XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;  
KW vector; vaccine; chicken; poultry; immunisation; ds.  
XX Gallid herpesvirus 1.  
XX WO9508622-A1.  
XX 30-MAR-1995.  
XX 16-SEP-1994; 94WO-US010628.  
XX 24-SEP-1993; 93US-00126597.  
XX (SYTR ) SYNTRO CORP.  
XX Wild MA, Cochran MD;  
XX WPI; 1995-139591/18.  
XX P-PSDB; AAW00642.

Recombinant attenuated infectious laryngotracheitis virus - for use in  
vaccines to protect poultry from infection from the virus, also methods  
of distinguishing between vaccinated and naturally infected birds.

Example 1; Page 115-116; 177pp; English.

ORF8 reverse complement (AAT33513) consists of bases 11,150 to 10,617 on  
the reverse complement strand of the unique short region (see also  
AAT33504) of infectious laryngotracheitis virus (ILTV) genomic DNA. It  
codes for a protein (AAW00642) of approx. 19,470 mol.wt. of unknown  
function. (Updated on 16-OCT-2003 to standardise OS field)

SQ Sequence 534 BP; 115 A; 138 C; 133 G; 148 T; 0 U; 0 Other;

Query Match 31.6%; Score 343; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 5e-99;  
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGAACCGACACCGCGCAACGAAATTCAAACGTTACTGTCCGCTCGGGCAATGAG 803  
DB 534 TTACGAACCGACACCGCGCAACGAAATTCAAACGTTACTGTCCGCTCGGGCAATGAG 475  
QY 804 CCTAGCTAATTTGGGTAAACCGTGGCTGCTGAGCGCAACGATCGGCTCGTCTCAT 863  
DB 474 CCTAGCTAATTTGGGTAAACCGTGGCTGCTGAGCGCAACGATCGGCTCGTCTCAT 415  
QY 864 TGTAAATTTCCATCGTCACCAAGAAACATGTGACCCCGCACCGAAATTAGACACGTTTC 923  
DB 414 TGTAAATTTCCATCGTCACCAAGAAACATGTGACCCCGCACCGAAATTAGACACGTTTC 355  
QY 924 GCAAGACGACGAAGACGTTTCCAAACTAGACGGGAATCGGAAATTTGGACCCATGTT 983

Db 354 GCAAGACGACGAGAACGCTTCCAACTAGAGGGAATCGGGAATAATTGGACCCATGGT 295  
QY 984 TCGGTGCGAAATAAACAAGGGGCGCTGACAGGATAGTGAACCTGTGGAACCTGGTTGCGAT 1043  
Db 294 TCGGTGCGAAATAAACAAGGGGCGCTGACAGGATAGTGAACCTGTGGAACCTGGTTGCGAT 235  
QY 1044 TGTAAACCGCTGCGCTGAAGTCCGCCGACTCAATAAAATG 1086  
Db 234 TGTAAACCGCTGCGCTGAAGTCCGCCGACTCAATAAAATG 192  
RESULT 6  
ADA71938/c  
ID ADA71938 standard; DNA; 2000 BP.  
XX  
AC ADA71938;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Rice gene, SEQ ID 5263.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX  
OS Oryza sativa.  
XX  
FN WC2003000898-A1.  
XX  
XX 03-JAN-2003.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
XX WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX  
XX Claim 27; SEQ ID NO 5263; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
XX  
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
SQ  
Query Match 4.4%; Score 47.8; DB 7; Length 2000;  
Best Local Similarity 7.5%; Pred. No. 0.00035;  
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;  
QY 390 AATCTCTGCGACGACGAGGGGATGTTCTAGTAATGTCGGCTAGACACCAAA 449  
Db 854 WWWWRYTMYTCYAMTCAKCKVAMTKMTTTCACARATSWRWRAMGRWRYKMKRA 795  
QY 450 AGAACCCATTGACGCTCTCCGCTCCCACTATCGGTGTATCAATTGCGAACACGCCGC 509  
Db 794 YWWRWFCWAGWAGWMMKSYRWKWKYATRYTMYKMWMTWWSWRWRSYRWWSGRM 735

QY 510 GACTCGGCGACTCTATTCCAAAGGCTGCTGTCGACACTTCGGGATTACCTACCGTCCAACT 569  
Db 734 RWSAWRYCGRMKCAKTKYASSARWTKRAKRSYRDRWYKRWKGTYRKYRWSRMTTRA 675  
QY 570 TGAGGCTCTACTCAGGACGAGGAAAGTTGGCGCAACTGCGCAAGCGTACTGTCACCGA 629  
Db 574 RSKRRKRWAGASMKSCWYWRGARSWMYKYSKSAKCKKTRTWTSSYMTSGWGYSSY 615  
QY 630 GGCCACGACGACGAGCGGCGGAGCGACAACCCGCGACGCGCGTCACTGCAACACGCGCTC 689  
Db 614 KSNWTSKNSYMGKMTCTMYTSMKGSRRSRKMGWSGMSRWYRWKGRKRYMYMKW 555  
QY 590 CGAATTGAAGCGGACACTTTTACCTTCCCTGGCTAGAAAATGGCGTGATCATACGA 749  
Db 554 KCTWRRCWYRWGYTMYTTSRSMYTGKRYKARYTSKRRYMYTKYKRYCWYTYGMYMKC 495  
QY 750 ACCGACACCGCGCAACGAAAATTCAAACGTTACTGTCGCTCGGACCAATGAGCCCTAC 809  
Db 494 SYMRVGYKACKCCCYAMCWKAAYSGMMYWKYSKWRMSTKYMWSWYKCKRSMKY 435  
QY 810 GCTAATTGGGGTAAACCGTGGCTGCTGCGAGCGCAACGATCGGCTCGCTCATTAAT 869  
Db 434 GAGCYCKMTYCSYGMKWTYMGSYKYSRCYKYMRYMYKGMWYMYSAYSMMTW 375  
QY 870 TTCATCGTCACGAGAAACATGTGACCCCGCGACCGGAAAATTAGACACGCTCTCG 924  
Db 374 YYYVAKYKWKYKRRGTMSWYKSKYKCYCTWNCYMKMERCYRWKMKRKKTKYS 320  
RESULT 7  
ADA71938  
ID ADA71938 standard; DNA; 2000 BP.  
XX  
AC ADA71938;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Rice gene, SEQ ID 5263.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX  
OS Oryza sativa.  
XX  
FN WC2003000898-A1.  
XX  
XX 03-JAN-2003.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
XX WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX  
XX Claim 27; SEQ ID NO 5263; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
XX  
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
SQ  
Query Match 4.4%; Score 47.8; DB 7; Length 2000;  
Best Local Similarity 7.5%; Pred. No. 0.00035;  
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;  
QY 390 AATCTCTGCGACGACGAGGGGATGTTCTAGTAATGTCGGCTAGACACCAAA 449  
Db 854 WWWWRYTMYTCYAMTCAKCKVAMTKMTTTCACARATSWRWRAMGRWRYKMKRA 795  
QY 450 AGAACCCATTGACGCTCTCCGCTCCCACTATCGGTGTATCAATTGCGAACACGCCGC 509  
Db 794 YWWRWFCWAGWAGWMMKSYRWKWKYATRYTMYKMWMTWWSWRWRSYRWWSGRM 735





Db 18034 TCCACGGCCACGCTGGCACCCTCGGTGGCGGAGCAACGGTGATCACCCCGTACGG 17975

Qy 89 TCTCCGCCAGGATTGACAGCATCATCGTGATCGTGGCGCTCGCCCGGAAGC 143

Db 17974 TCGGCCCCAGGACGGTGTCTCCAGCGGAGCGGGAACCGCGCAAGCACCGCGGC 17920

RESULT 11

AAI199683\_06/c

Continuation (7 of 44) of AAI199683 from base 600001 (Mycobacterium tuberculosis strain H

WP Sequence split into 44 fragments LOCUS AAI199683 Accession Aai199683

Fragment Name Begin End

WP AAI199683\_00 1 110000

WP AAI199683\_01 100001 210000

WP AAI199683\_02 200001 310000

WP AAI199683\_03 300001 410000

WP AAI199683\_04 400001 510000

WP AAI199683\_05 500001 610000

WP AAI199683\_06 600001 710000

WP AAI199683\_07 700001 810000

WP AAI199683\_08 800001 910000

WP AAI199683\_09 900001 1010000

WP AAI199683\_10 1000001 1110000

WP AAI199683\_11 1100001 1210000

WP AAI199683\_12 1200001 1310000

WP AAI199683\_13 1300001 1410000

WP AAI199683\_14 1400001 1510000

WP AAI199683\_15 1500001 1610000

WP AAI199683\_16 1600001 1710000

WP AAI199683\_17 1700001 1810000

WP AAI199683\_18 1800001 1910000

WP AAI199683\_19 1900001 2010000

WP AAI199683\_20 2000001 2110000

WP AAI199683\_21 2100001 2210000

WP AAI199683\_22 2200001 2310000

WP AAI199683\_23 2300001 2410000

WP AAI199683\_24 2400001 2510000

WP AAI199683\_25 2500001 2610000

WP AAI199683\_26 2600001 2710000

WP AAI199683\_27 2700001 2810000

WP AAI199683\_28 2800001 2910000

WP AAI199683\_29 2900001 3010000

WP AAI199683\_30 3000001 3110000

WP AAI199683\_31 3100001 3210000

WP AAI199683\_32 3200001 3310000

WP AAI199683\_33 3300001 3410000

WP AAI199683\_34 3400001 3510000

WP AAI199683\_35 3500001 3610000

WP AAI199683\_36 3600001 3710000

WP AAI199683\_37 3700001 3810000

WP AAI199683\_38 3800001 3910000

WP AAI199683\_39 3900001 4010000

WP AAI199683\_40 4000001 4110000

WP AAI199683\_41 4100001 4210000

WP AAI199683\_42 4200001 4310000

WP AAI199683\_43 4300001 4403765

Query Match 3.5%; Score 38.2; DB 4; Length 110000;

Best Local Similarity 58.3%; Pred. No. 2.9; Mismatches 48; Indels 0; Gaps 0;

Matches 67; Conservative 0

Qy 29 TCTTTCGCCGCGACGCTCGCACCCCTTCGGCGCGATGGGAATCGTGATCGTACATCGAATACAG 88

Db 19394 TCCACGCGCCACGCTGGCACCTCGGTGGCGGAGCAACGGTGATCACCCCGGTACGG 19335

Qy 89 TCTCCGCCAGGATTGACAGCATCATCGTGATCGTGGCGCTCGCCCGGAAGC 143

Db 19334 TCGGCCCCAGGACGGTGTCCAGCGGAGCGGGAACACCGCGCAAGCACCGCGGC 19280

RESULT 12

AAF26455

ID AAF26455 standard; cDNA; 2526 BP.

XX AAF26455;

AC 26-MAR-2001 (first entry)

XX D.melanogaster corkscrew DNA.

DE

XX SHP-2; SHP-1; Src Homology-2; protein tyrosine phosphatase; mutant;

KW neoplastic disorder; obesity; angiogenesis; cancer; immune;

KW hematopoietic; allergy; ss.

XX Drosophila melanogaster.

XX US6156551-A.

PN 05-DEC-2000.

XX 05-JUN-1998; 98US-00092443.

XX 05-JUN-1998; 98US-00092443.

PR (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PA (JOSL-) JOSLIN DIABETES CENT.

XX Neel BG, Shoelson S, Pluskey S, O'reilly AM;

XX WPI; 2001-060166/07.

PT Mutant SH2 domain-containing protein tyrosine phosphatase, useful in in

PT vitro assays to screen for binding partners, inhibitors of tyrosine

XX phosphatase and for treating tyrosine phosphatase-mediated diseases.

PS Disclosure; Fig 7; 161pp; English.

XX The present invention relates to an activated SH2 (Src Homology-2) -

CC domain containing protein tyrosine phosphatase (SHP-1 or SHP-2) mutant

CC with a mutation in the SH2-domain. Activated mutants of SH2-domain-

CC containing protein tyrosine phosphatases are useful in in vitro assays to

CC screen for binding partners and inhibitors of the phosphatase and in the

CC treatment of PTP-mediated diseases or conditions in a mammal, including

CC neoplastic disorders, obesity and to inhibit angiogenesis. Inhibitors

CC identified using the activated mutants are useful for the treatment of

CC cancer, immunosuppression, immunostimulation, hematopoietic stimulation

CC and anti-allergy treatment

XX SQ Sequence 2526 BP; 616 A; 742 C; 732 G; 436 T; 0 U; 0 Other;

Query Match 3.5%; Score 38; DB 5; Length 2526;

Best Local Similarity 51.8%; Pred. No. 0.53;

Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 569 TTGAGGCTATCTCAGACCGAGGAAGTTGGCGCACTGGCAAGCGTACGTTGCCACGG 628

Db 1325 TTAAGACCTACATCGCCACCCAGGGTGTCTGTCTACCCAGCAAGTGAACACCGTGACGG 1384

Qy 629 AGGCCAGCAGCAGCGCGCGGAGCAACCCCGACGCCCGCTCACTGCAACACCGAGCCT 688

Db 1385 ACTTCTGGAACATGGTCTGCGCAGGAACACCGCGGTGATCTGTCATGACCAACAGAGT 1444

Qy 689 CCGAACCTGAAGCGGAACTTATCTTCCCTGGGTAGAAATGG 734

Db 1445 ACAGCGCGCAAGAAAGTGGCGCGCTACTTGGCGGACGAGGG 1490

RESULT 13

ABL06927

ID ABL06927 standard; cDNA; 2677 BP.

XX ABL06927;

AC

XX

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 15263.

```
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 15263; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2677 BP; 670 A; 760 C; 749 G; 498 T; 0 U; 0 Other;
SQ
Query Match 3.5%; Score 38; DB 4; Length 2677;
Best Local Similarity 51.8%; Pred. No. 0.55;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 569 TTGAGGCCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
Db 1464 TTAAGACCTACATCGCCACCCAGGGCTGTCTGCTCACCAGCAAGTGAACACGGTGACGG 1523
QY 629 AGGCCAGCAGCAGCGCCGAGCGGACACACCCGAGCCCGCTCACTGCAACCGCGCCT 688
Db 1524 ACTTCTGGAACATGCTGTGGCAGGAGAACACGCGGGTGTATCGTCATGACCAACCAAGGAGT 1583
QY 689 CCGNACTTGAAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
Db 1584 ACGAGCGCGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1629
RESULT 14
ABL18521
ID ABL18521 standard; DNA; 2838 BP.
XX ABL18521;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7036.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 7036; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2838 BP; 711 A; 853 C; 799 G; 475 T; 0 U; 0 Other;
SQ
Query Match 3.5%; Score 38; DB 4; Length 2838;
Best Local Similarity 51.8%; Pred. No. 0.57;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 569 TTGAGGCCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
Db 1625 TTAAGACCTACATCGCCACCCAGGGCTGTCTGCTCACCAGCAAGTGAACACCGTGACGG 1684
QY 629 AGGCCAGCAGCAGCGCCGAGCGGACACACCCGAGCCCGCTCACTGCAACCGCGCCT 688
Db 1685 ACTTCTGGAACATGCTGTGGCAGGAGAACACGCGGGTGTATCGTCATGACCAACCAAGGAGT 1744
QY 689 CCGNACTTGAAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
Db 1745 ACGAGCGCGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1790
RESULT 15
ABL21135
ID ABL21135 standard; DNA; 4528 BP.
XX ABL21135;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 14878.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
```

PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
XX  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX  
XX Claim 1; SEQ ID NO 14878; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABBS72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 4528 BP; 1311 A; 1118 C; 1051 G; 1048 T; 0 U; 0 Other;  
Query Match 3.5%; Score 38; DB 4; Length 4528;  
Best Local Similarity 51.8%; Pred. No. 0.71;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 569 TTGAGGCTATCTCAGGACCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628  
Db 1464 TTAAGACCTACATCGCCACCCAGGGCTGTCTGCTCACCAGCAAGTGAACACGGTGACGG 1523  
QY 629 AGGCCACGACGACCGCGCGGCGGCAACACCCGACGCCGCTCACTGCAACCGCGGCT 688  
Db 1524 ACTTCTGGAACATGCTCTGGCAGGAGAACACCGCGGTGATGTCATGACCCACCGAGAGT 1583  
QY 689 CCGAACTTGAACGGAACACTTTACCTTTCCCTGGCTAGAAAAGG 734  
Db 1584 ACGAGCGCGCAAGAAAAGTCCGCCCTACTGCGCGGACGAGGG 1629

Search completed: March 7, 2004, 23:09:33  
Job time : 556 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 22:43:59 ; Search time 120 Seconds  
(without alignments)  
5022.306 Million cell updates/sec

Title: US-09-993-777-1\_COPY\_9874\_10959

Perfect score: 1086  
Sequence: 1 ATGGCATCGTACTTGGAC.....CGCCGACTCAATAAAATG 1086

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	13473	5	PCT-US96-03916-1
2	1084.4	99.9	18912	5	PCT-US96-03916-59
3	343	31.6	534	5	PCT-US96-03916-18
4	38.2	3.5	4403765	3	US-09-103-840A-2
5	38.2	3.5	4411529	3	US-09-103-840A-1
6	35.6	3.3	420	4	US-09-252-991A-4761
7	35.6	3.3	642	4	US-09-252-991A-4683
8	35.6	3.3	960	4	US-09-252-991A-4705
9	35.6	3.3	1554	4	US-09-252-991A-4723
10	35.2	3.2	19718	4	US-08-961-527-99
11	34.6	3.2	50341	1	US-08-247-901C-1
12	34.6	3.2	50341	2	US-09-075-904-1
13	34.6	3.2	52297	3	US-09-426-436-1
14	34.6	3.2	52297	4	US-08-705-557-1
15	33.2	3.1	1230025	4	US-09-198-452A-1
16	33	3.0	1068	4	US-09-489-039A-689
17	32.6	3.0	2010	4	US-09-527-972-3
18	32.4	3.0	3024	4	US-09-833-381-889
19	32.2	3.0	4403765	3	US-09-103-840A-2
20	32.2	3.0	4411529	3	US-09-103-840A-1
21	31.8	2.9	1145	3	US-08-809-103B-7
22	31.8	2.9	1148	3	US-08-809-103B-1
23	31.8	2.9	1150	3	US-08-809-103B-3
24	31.8	2.9	1150	3	US-08-809-103B-5
25	31.8	2.9	7218	1	US-08-232-463-14
26	31.6	2.9	843	3	US-08-513-974B-375
27	31.6	2.9	25165	4	US-09-453-702B-39

C	28	30.8	2.8	948	4	US-09-134-001C-1418	Sequence 1418, Ap
	29	30.8	2.8	1728	1	US-08-427-097-1	Sequence 1, Appl
	30	30.8	2.8	1728	2	US-08-878-957-1	Sequence 1, Appl
	31	30.8	2.8	1746	4	US-09-489-039A-4325	Sequence 4325, Ap
	32	30.8	2.8	1752	1	US-08-427-097-15	Sequence 15, Appl
	33	30.8	2.8	1752	1	US-08-427-097-27	Sequence 27, Appl
	34	30.8	2.8	1752	2	US-08-878-957-15	Sequence 15, Appl
	35	30.8	2.8	1752	2	US-08-878-957-27	Sequence 27, Appl
C	36	30.6	2.8	245	4	US-09-313-294A-3607	Sequence 3607, Ap
	37	30.6	2.8	505	4	US-09-621-976-15639	Sequence 15639, A
C	38	30.6	2.8	795	4	US-09-252-991A-4788	Sequence 4788, Ap
	39	30.6	2.8	1047	4	US-09-252-991A-14678	Sequence 14678, A
	40	30.6	2.8	1512	4	US-09-252-991A-1320	Sequence 1320, Ap
	41	30.6	2.8	1614	4	US-09-616-289-45	Sequence 45, Appl
	42	30.6	2.8	1686	4	US-09-252-991A-14548	Sequence 14548, A
C	43	30.6	2.8	1815	4	US-09-252-991A-15077	Sequence 15077, A
C	44	30.6	2.8	2130	4	US-09-252-991A-1148	Sequence 1148, Ap
	45	30.6	2.8	12425	4	US-09-616-289-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1  
PCT-US96-03916-1  
Sequence 1, Application PC/TUS9603916  
GENERAL INFORMATION:  
APPLICANT: Wild, Martha A.  
APPLICANT: Cochran, Mark D.  
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/03916  
FILING DATE: 23-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/126,597  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 39116-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13473 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1059..2489  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2575..4107  
FEATURE:

```
NAME/KEY: CDS
LOCATION: 4113..4445
FEATURE:
NAME/KEY: CDS
LOCATION: 4609..5487
FEATURE:
NAME/KEY: CDS
LOCATION: 5697..8654
FEATURE:
NAME/KEY: CDS
LOCATION: 9874..10962
FEATURE:
NAME/KEY: CDS
LOCATION: 11159..12658
FEATURE:
NAME/KEY: CDS
LOCATION: 12665..13447
PCT-US96-03916-1

Query Match      100.0%; Score 1086; DB 5; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCATCGCTACTTGGAACTCTGCTCTCTCTTGGCGGAGCGTCGACCCCTTCGGGCGG 60
DB 9874 ATGGCATCGCTACTTGGAACTCTGCTCTCTCTTGGCGGAGCGTCGACCCCTTCGGGCGG 9933
QY 61 ATGGGAATCGTGATCACTGGAAATCACTCTCCGCGCAGGATGAGCAGATCATCATCGTG 120
DB 9934 ATGGGAATCGTGATCACTGGAAATCACTCTCCGCGCAGGATGAGCAGATCATCATCGTG 9993
QY 121 ATCTCGCGGCTCGCGCGGAGCTACAACTCACTGAGCTATTTTCATGCTGCGGCGAG 180
DB 9994 ATCTCGCGGCTCGCGCGGAGCTACAACTCACTGAGCTATTTTCATGCTGCGGCGAG 10053
QY 181 AGACCCCAACAACCTACTCAGGAACCGTCGCGTTCGGTTCGGTTCGATATACAAAC 240
DB 10054 AGACCCCAACAACCTACTCAGGAACCGTCGCGTTCGGTTCGGTTCGATATACAAAC 10113
QY 241 CAGTGCTACCGAGAACTTAGCGAGGAGCGCTTTGAAATTCACATCATCGATCGTCTCT 300
DB 10114 CAGTGCTACCGAGAACTTAGCGAGGAGCGCTTTGAAATTCACATCATCGATCGTCTCT 10173
QY 301 GTTTTGTGCGGTGTAAGTAGTACCGAGTACAGTTCTCGGCTCGAAGAGTAAACCGGA 360
DB 10174 GTTTTGTGCGGTGTAAGTAGTACCGAGTACAGTTCTCGGCTCGAAGAGTAAACCGGA 10233
QY 361 CTTCCACACCGGTTTAAAGCTCACTATAGAAATCTCTCGTTCGGAACGACGAGGATGTC 420
DB 10234 CTTCCACACCGGTTTAAAGCTCACTATAGAAATCTCTCGTTCGGAACGACGAGGATGTC 10293
QY 421 TAGCTAATTTTTCGGCTAGACACCAAGAACCACTTGACGCTTCGCGATCCAACTA 480
DB 10294 TAGCTAATTTTTCGGCTAGACACCAAGAACCACTTGACGCTTCGCGATCCAACTA 10353
QY 481 TCGGTGTATCAATTCGGAACACCGCGCGACTTCGCGGACTCTATTCGAAGGCTTCGTGT 540
DB 10354 TCGGTGTATCAATTCGGAACACCGCGCGACTTCGCGGACTCTATTCGAAGGCTTCGTGT 10413
QY 541 CGCACCTTCGGATTACTACCGTCCAACTTAGGCGCTATCTCAGAACCGAGGAAGTTGG 600
DB 10414 CGCACCTTCGGATTACTACCGTCCAACTTAGGCGCTATCTCAGAACCGAGGAAGTTGG 10473
QY 601 CGCACTGCGAAGCGTACTGTCGAGGAGGCGCAGACGACGAGCGGCGAGGCGACACAC 660
DB 10474 CGCACTGCGAAGCGTACTGTCGAGGAGGCGCAGACGACGAGCGGCGAGGCGACACAC 10533
QY 661 CGGACGCGGCTCACTGCAACGAGCGCTCCGAATTTGAAGCGGAACATTTACCTTTCCC 720
DB 10534 CGGACGCGGCTCACTGCAACGAGCGCTCCGAATTTGAAGCGGAACATTTACCTTTCCC 10593
QY 721 TGGCTAGAAATGGCGTGGATCATTAAGAACCGACCGGCAACGAAATTCAAACGTT 780
```

```
DB 10594 TGGCTAGAAATGGCGTGGATCATTAAGAACCGACCGGCAACGAAATTCAAACGTT 10653
QY 781 ACTGTCGCTTCGGGACCAATGAGCCCTACGCTAAATTTGGGTAAACCGTGGCTGCTCGTG 840
DB 10654 ACTGTCGCTTCGGGACCAATGAGCCCTACGCTAAATTTGGGTAAACCGTGGCTGCTCGTG 10713
QY 841 AGCGCAACGATCGGCTCGTCAATTTGTAATTTCCATGCTCAACGAAACATGTCGACCCCG 900
DB 10714 AGCGCAACGATCGGCTCGTCAATTTGTAATTTCCATGCTCAACGAAACATGTCGACCCCG 10773
QY 901 CACCGAAATTTAGACACCGTCTCGCAAGACGAGAACGTTCCCAACTAGAGGGAA 960
DB 10774 CACCGAAATTTAGACACCGTCTCGCAAGACGAGAACGTTCCCAACTAGAGGGAA 10833
QY 961 TCGGCAAAATTTGGACCCCATGTTGCGTCCGAAATAAACAAGGGCGCTGACGAGTACT 1020
DB 10834 TCGGCAAAATTTGGACCCCATGTTGCGTCCGAAATAAACAAGGGCGCTGACGAGTACT 10893
QY 1021 GAATTTGTGGAACCTGGTTCGATTTGTAACCGTCTGGCTAAGCTCGCCCGACTCAATA 1080
DB 10894 GAATTTGTGGAACCTGGTTCGATTTGTAACCGTCTGGCTAAGCTCGCCCGACTCAATA 10953
QY 1081 AAAATG 1086
DB 10954 AAAATG 10959

RESULT 2
PCT-US96-03916-59
Sequence 59, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 18912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
```

```
LOCATION: 697..1533
FEATURE:
NAME/KEY: CDS
LOCATION: complement (1900..2784)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: complement (2916..3605)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 3694..5124
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 5210..7081
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 7245..8123
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 8333..11290
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 11098..12402
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 12510..13598
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 13792..15291
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 15298..16080
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 16129..17013
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: complement (17380..18216)
OTHER INFORMATION:
PCT-US96-03916-59

Query Match
Best Local Similarity 99.9%; Score 1084.4; DB 5; Length 18912;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCATCGTACTTGGAACTCTGGCTCTCTCTTGGCGGAGCGCTCGCACCCCTTCGCGCG 60
DDB 12510 ATGGCATCGTACTTGGAACTCTGGCTCTCTCTTGGCGGAGCGCTCGCACCCCTTCGCGCG 12569

QY 61 ATGGGAATCGTATCACTGGAATCACTCTCCGCGAGGATTCAGCAGCATCACATCGTG 120
DDB 12570 ATGGGAATCGTATCACTGGAATCACTCTCCGCGAGGATTCAGCAGCATCACATCGTG 12629

QY 121 ATGCTCGCGCTCGCCCGGAGCTACAATTCAGCTCAATTCAGCTATTTTCATGCTGCCAG 180
DDB 12630 ATGCTCGCGCTCGCCCGGAGCTACAATTCAGCTCAATTCAGCTATTTTCATGCTGCCAG 12689

QY 181 AGACCCCAAAACCTACTCAGAACCGTCCGGTTCGGTTCGGTCTGATATAACAAC 240
DDB 12690 AGACCCCAAAACCTACTCAGAACCGTCCGGTTCGGTTCGGTCTGATATAACAAC 12749

QY 241 CAGTGCTACAGGAACTTAGCAGGAGCGCTTTGAAAATTGCACTCATCGATCTTCT 300
```

```
Db 12750 CAGTGCTACAGGAACTTAGCAGGAGCGCTTTGAAAATTGCACTCATCGATCTTCT 12809
QY 301 GTTTTTCGCGCTTAAAGTGACCGAGTACAGCGTTCTCCGCTCGAACAGACTAACCGGA 360
DDB 12810 GTTTTTCGCGCTTAAAGTGACCGAGTACAGCGTTCTCCGCTCGAACAGACTAACCGGA 12869
QY 361 CCTCCACACCCGTTTAAAGTCACTATACGAAATCCTCGTCCGAAACGACAGCGGATGTC 420
DDB 12870 CCTCCACACCCGTTTAAAGTCACTATACGAAATCCTCGTCCGAAACGACAGCGGATGTC 12929
QY 421 TACGTAATTTGTTTCGGCTAGACGACACCAAGAACCCATTTGAGCTTTCGCGATCAACTA 480
DDB 12930 TACGTAATTTGTTTCGGCTAGACGACACCAAGAACCCATTTGAGCTTTCGCGATCAACTA 12989
QY 481 TCGGTGTATCAATTCGCGAACACCGCGGACTCGGAGCTCTATTCCAGGCTTCGTGT 540
DDB 12990 TCGGTGTATCAATTCGCGAACACCGCGGACTCGGAGCTCTATTCCAGGCTTCGTGT 13049
QY 541 CGCACTTCGGATTACCTTACCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAAGTTGG 600
DDB 13050 CGCACTTCGGATTACCTTACCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAAGTTGG 13109
QY 601 CGCACTTCGGCAAGCGGTACGTTGCCACGAGGAGCCACGACGACGCGCGGAGCGACAACC 660
DDB 13110 CGCACTTCGGCAAGCGGTACGTTGCCACGAGGAGCCACGACGACGCGCGGAGCGACAACC 13169
QY 661 CCGACGCGCGTCACTGCAACACGAGCGCTTCGAACTTGAAGCGGAACTTTTACCTTTCCC 720
DDB 13170 CCGACGCGCGTCACTGCAACACGAGCGCTTCGAACTTGAAGCGGAACTTTTACCTTTCCC 13229
QY 721 TGGCTAGAAAATGCGGTGGATCATAGAACGAGCACCGCAACGAAATTCGAAACGTT 780
DDB 13230 TGGCTAGAAAATGCGGTGGATCATAGAACGAGCACCGCAACGAAATTCGAAACGTT 13289
QY 781 ACTGTCGCTTCGGGACAAATGAGCCCTACGCTAATTTGGGGTAACGCTGGCTCGCTCGTG 840
DDB 13290 ACTGTCGCTTCGGGACAAATGAGCCCTACGCTAATTTGGGGTAACGCTGGCTCGCTCGTG 13349
QY 841 AGCGCAACGATCGGCTCGCTCATTTGTAATTTCCATCGTCACAGAAACATGTGACCCCG 900
DDB 13350 AGCGCAACGATCGGCTCGCTCATTTGTAATTTCCATCGTCACAGAAACATGTGACCCCG 13409
QY 901 CACGAAATTTAGACAGGTTCTCGCAAGACGACGAGAAACGTTCCGAAACTAGAGGGAA 960
DDB 13410 CACGAAATTTAGACAGGTTCTCGCAAGACGACGAGAAACGTTCCGAAACTAGAGGGAA 13469
QY 961 TCGGAAAATTTGGACCCATGTTGCGTGGGAAATACAGGCGGCTGACAGGATAGT 1020
DDB 13470 TCGGAAAATTTGGACCCATGTTGCGTGGGAAATACAGGCGGCTGACAGGATAGT 13529
QY 1021 GAACTTGTGAACTGGTTGCGATTGTTAACCCGCTGCGCTAAGCTCGCCGACTCAATA 1080
DDB 13530 GAACTTGTGAACTGGTTGCGATTGTTAACCCGCTGCGCTAAGCTCGCCGACTCAATA 13589
QY 1081 AAAATG 1086
DDB 13590 AAAATG 13595
```

## RESULT 3

PCT-US96-03916-18/c

; Sequence 18, Application PC/TUS9603916

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESS: Cooper &amp; Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/03916  
FILING DATE: 23-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/136,597  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 39116-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 534 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..533  
PCT-US96-03916-18

Query Match 31.6%; Score 343; DB 5; Length 534;  
Best Local Similarity 100.0%; Pred. No. 4.9e-101;  
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAACCAAAATTCACGTTACTGTCGGTCTCGGACCAATGAG 803  
DB 534 TTACGACCGACACCGCGCAACCAAAATTCACGTTACTGTCGGTCTCGGACCAATGAG 475  
QY 804 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGTGTGAGCGCAACGATCGGCTCGTCAT 863  
DB 474 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGTGTGAGCGCAACGATCGGCTCGTCAT 415  
QY 864 TGTAATTTCCATCGTCACAGAAACATGTGCACCCCGCACCGAAATTTAGACACGCTCTC 923  
DB 414 TGTAATTTCCATCGTCACAGAAACATGTGCACCCCGCACCGAAATTTAGACACGCTCTC 355  
QY 924 GCAAGACGACGAGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCCATGTT 983  
DB 354 GCAAGACGACGAGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCCATGTT 295  
QY 984 TGGGTGCGAATAAACAAGGCGCTGACGAGGATAGTGAATTTGGAATGTTGGCAT 1043  
DB 294 TGGGTGCGAATAAACAAGGCGCTGACGAGGATAGTGAATTTGGAATGTTGGCAT 235  
QY 1044 TGTTAAACCGCTGTCGGCTAGCTCGCCCGACTCAATAAAAATG 1086  
DB 234 TGTTAAACCGCTGTCGGCTAGCTCGCCCGACTCAATAAAAATG 192

RESULT 4  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.

STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/03916  
FILING DATE: 23-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/136,597  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 39116-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 534 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..533  
PCT-US96-03916-18

Query Match 31.6%; Score 343; DB 5; Length 534;  
Best Local Similarity 100.0%; Pred. No. 4.9e-101;  
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAACCAAAATTCACGTTACTGTCGGTCTCGGACCAATGAG 803  
DB 534 TTACGACCGACACCGCGCAACCAAAATTCACGTTACTGTCGGTCTCGGACCAATGAG 475  
QY 804 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGTGTGAGCGCAACGATCGGCTCGTCAT 863  
DB 474 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGTGTGAGCGCAACGATCGGCTCGTCAT 415  
QY 864 TGTAATTTCCATCGTCACAGAAACATGTGCACCCCGCACCGAAATTTAGACACGCTCTC 923  
DB 414 TGTAATTTCCATCGTCACAGAAACATGTGCACCCCGCACCGAAATTTAGACACGCTCTC 355  
QY 924 GCAAGACGACGAGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCCATGTT 983  
DB 354 GCAAGACGACGAGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCCATGTT 295  
QY 984 TGGGTGCGAATAAACAAGGCGCTGACGAGGATAGTGAATTTGGAATGTTGGCAT 1043  
DB 294 TGGGTGCGAATAAACAAGGCGCTGACGAGGATAGTGAATTTGGAATGTTGGCAT 235  
QY 1044 TGTTAAACCGCTGTCGGCTAGCTCGCCCGACTCAATAAAAATG 1086  
DB 234 TGTTAAACCGCTGTCGGCTAGCTCGCCCGACTCAATAAAAATG 192

RESULT 4  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 3.5%; Score 38.2; DB 3; Length 4403765;  
Best Local Similarity 58.3%; Pred. No. 3.6;  
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCCCTTCGGCGGATGGGAATCGTGATCCTGGAATCAGC 88  
DB 619394 TCCACGCGCCACGCTGCGCACCCCTTCGGTGGCGGACGCGTGATCACCCTCGTCAGCG 619335  
QY 89 TCTCCGCGCAGATTGACGACGATCAGATCGTGATCGTGGCGCTCGCCCGAAGC 143  
DB 619334 TCGGCCCGGACGCGGTGTCCAGCGGACGACCGCGCCCAAGCAGCAGCGCGC 619280

RESULT 5  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1

Query Match 3.5%; Score 38.2; DB 3; Length 4411529;  
Best Local Similarity 58.3%; Pred. No. 3.6;  
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCCCTTCGGCGGATGGGAATCGTGATCCTGGAATCAGC 88  
DB 618034 TCCACGCGCCACGCTGCGCACCCCTTCGGTGGCGGACGCGTGATCACCCTCGTCAGCG 617975  
QY 89 TCTCCGCGCAGATTGACGACGATCAGATCGTGATCGTGGCGCTCGCCCGAAGC 143  
DB 617974 TCGGCCCGGACGCGGTGTCCAGCGGACGACCGCGCCCAAGCAGCAGCGCGC 617920

RESULT 6  
US-09-252-991A-4761/c  
; Sequence 4761, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: MARC J. RUBENFIELD ET AL.

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 4761  
;; LENGTH: 420  
;; TYPE: DNA  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4761

Query Match 3.3%; Score 35.6; DB 4; Length 420;  
Best Local Similarity 51.2%; Pred. No. 0.16;  
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 508 GCGACTCGGAGCTATTTCGAAGCTTCGTGTCGCACCTTCGGATTACCTACCGTCCAA 567  
Db 276 GCGCGGCGCGCGCTGCGCACTGTTCGTCTCGGCCATCCGCGACCGACCGGCAAC 217  
QY 568 CTTGAGGCTATCTCAGGACGAGGAAAGTTGGCGCAACTGGCAAGCTAGTTCACG 627  
Db 216 GCGGCGCGCGCTTGACCGCGCGACGAGGCGGCGTCTGTGAGGAGTTCGCGCGCAGG 157  
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669  
Db 156 GTGGCGCCAGCGAGTACTCTCGAGGACGCGACCTGCGCGCGC 115

## RESULT 7

US-09-252-991A-4683  
; Sequence 4683, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 4683  
;; LENGTH: 642  
;; TYPE: DNA  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4683

Query Match 3.3%; Score 35.6; DB 4; Length 642;  
Best Local Similarity 51.2%; Pred. No. 0.2;  
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 508 GCGACTCGGAGCTATTTCGAAGCTTCGTGTCGCACCTTCGGATTACCTACCGTCCAA 567  
Db 280 GCGCGGCGCGCGCTGCGCACTGTTCGTCTCGGCCATCCGCGACCGACCGGCAAC 339  
QY 568 CTTGAGGCTATCTCAGGACGAGGAAAGTTGGCGCAACTGGCAAGCTAGTTCACG 627  
Db 340 GCGGCGCGCGCTTGACCGCGCGACGAGGCGCGCTCTGAGGACGTCGCGCGCAGG 399  
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669  
Db 400 GTGGCGCCAGCGAGTACTCTGAGGACGCGCGACCTGCGCGCGC 441

## RESULT 8

US-09-252-991A-4705  
; Sequence 4705, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 4705  
;; LENGTH: 960  
;; TYPE: DNA  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4705

Query Match 3.3%; Score 35.6; DB 4; Length 960;  
Best Local Similarity 51.2%; Pred. No. 0.25; 79; Indels 0; Gaps 0;  
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 508 GCGACTCGGAGCTATTTCGAAGCTTCGTGTCGCACCTTCGGATTACCTACCGTCCAA 567  
Db 213 GCGCGGCGCGCGCTGCGCACTGTTCGTCTCGGCCATCCGCGACCGACCGGCAAC 272  
QY 568 CTTGAGGCTATCTCAGGACGAGGAAAGTTGGCGCAACTGGCAAGCTAGTTCACG 627  
Db 273 GCGGCGCGCGCTTGACCGCGCGGCGGCGTCTGTGAGGACGTCGCGCGCAGG 332  
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669  
Db 333 GTGGCGCCAGCGAGTACTCTGAGGACGCGCGACCTGCGCGCGC 374

## RESULT 9

US-09-252-991A-4723  
; Sequence 4723, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 4723  
;; LENGTH: 1554  
;; TYPE: DNA  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4723

Query Match 3.3%; Score 35.6; DB 4; Length 1554;  
Best Local Similarity 51.2%; Pred. No. 0.33;  
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 508 GCGACTCGGAGCTATTTCGAAGCTTCGTGTCGCACCTTCGGATTACCTACCGTCCAA 567  
Db 1112 GCGCGGCGCGCGCTGCGCACTGTTCGTCTCGGCCATCCGCGACCGACCGGCAAC 1171  
QY 568 CTTGAGGCTATCTCAGGACGAGGAAAGTTGGCGCAACTGGCAAGCTAGTTCACG 627  
Db 1172 GCGGCGCGCGCTTGACCGCGCGGCGGCGTCTGTGAGGACGTCGCGCGCAGG 1231  
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669

DB 1232 GTGGCGCAGCGAGTACTCTGAGGACGCCGACCTGCGCGGC 1273

STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,901C  
FILING DATE: May 23, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/057,531  
FILING DATE: April 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Elizabeth A  
REGISTRATION NUMBER: 39,911  
REFERENCE/DOCKET NUMBER: 96700/273  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50341  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE:  
DESCRIPTION: L5 shuttle phasmid sequence  
HYPOTHETICAL: NO  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: L5 mycobacteriophage  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: No. 5750384e  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-247-901C-1

Query Match 3.2%; Score 34.6; DB 1; Length 50341;  
Best Local Similarity 49.2%; Pred. No. 4.6;  
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

431 TTCGGCTAGACACACCAAGAACCCATTGACCTTCGCGATCCCACTATCGGTATC 490

1232 GTGGCGCAGCGAGTACTCTGAGGACGCCGACCTGCGCGGC 1273

STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,901C  
FILING DATE: May 23, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/057,531  
FILING DATE: April 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Elizabeth A  
REGISTRATION NUMBER: 39,911  
REFERENCE/DOCKET NUMBER: 96700/273  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50341  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE:  
DESCRIPTION: L5 shuttle phasmid sequence  
HYPOTHETICAL: NO  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: L5 mycobacteriophage  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: No. 5750384e  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-247-901C-1

Query Match 3.2%; Score 34.6; DB 1; Length 50341;  
Best Local Similarity 49.2%; Pred. No. 4.6;  
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

431 TTCGGCTAGACACACCAAGAACCCATTGACCTTCGCGATCCCACTATCGGTATC 490

Db 18665 TTCGGCTGTTGAGGATGACGACCCCAAGACGATGCTCGCTTCAGGGCTCGGTGGAC 18606  
Qy 491 AATTGCGAACAACCGCCGCGACTTCGCGACTCTATTCCAAGGCTTCGTGCGACCTTCG 550  
Db 18605 AAGTAGCGAGCGCGGGTAGTTCGCCGCTCTTCAACAACGACCTTCGACGGAGGTG 18546  
Qy 551 GATTACCTACGTCGCACTTGAAGGCTTCTCAGGACGGAGAAAGTTGGCGCACTGGC 610  
Db 18545 TAGAACAACCCCTCCACGCTGTGTGGCCAGGTACACACCCCTGGTCACCGGTGCTCAAGTTG 18486  
Qy 611 AAGCG 615  
Db 18485 AAACG 18481

## RESULT 12

US-09-075-904-1/c  
; Sequence 1, Application US/09075904  
; Patent No. 5994137

## GENERAL INFORMATION:

APPLICANT: Jacobs, et al.  
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESS: Amster, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,904  
FILING DATE: May 11, 1998

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/247,901  
FILING DATE: May 23, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Elizabeth A  
REGISTRATION NUMBER: 39,911  
REFERENCE/DOCKET NUMBER: 96700/475  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 50341  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: L5 shuttle phasmid sequence  
HYPOTHETICAL: No  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: L5 mycobacteriophage  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:

## CHROMOSOME/SEGMENT:

FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: No. 5994137e  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-09-075-904-1

Query Match 3.2%; Score 34.6; DB 2; Length 50341;  
Best Local Similarity 49.2%; Pred. No. 4.6;  
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
Qy 431 TTCGGCTAGACGACACCAAGAACCCATTGACGTCTTCGCGATCCAACTATCGGTATC 490  
Db 18665 TTCGGCTGTTGAGGATGACGACCCCAAGACGATGCTCGCTTCAGGGCTCGGTGGAC 18606  
Qy 491 AATTGCGAACAACCGCCGCGACTTCGCGACTCTATTCCAAGGCTTCGTGCGACCTTCG 550  
Db 18605 AAGTAGCGAGCGCGGGTAGTTCGCCGCTCTTCAACAACGACCTTCGACGGAGGTG 18546  
Qy 551 GATTACCTACGTCGCACTTGAAGGCTTCTCAGGACGGAGAAAGTTGGCGCACTGGC 610  
Db 18545 TAGAACAACCCCTCCACGCTGTGTGGCCAGGTACACACCCCTGGTCACCGGTGCTCAAGTTG 18486  
Qy 611 AAGCG 615  
Db 18485 AAACG 18481

## RESULT 13

US-09-426-436-1/c  
; Sequence 1, Application US/09426436  
; Patent No. 6225066

## GENERAL INFORMATION:

APPLICANT: William R. Jacobs, Jr.  
APPLICANT: Barry R. Bloom  
APPLICANT: Graham F. Hatfull  
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC  
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESS: Amster, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/426,436  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,557  
FILING DATE:  
APPLICATION NUMBER: US/08/057,531  
FILING DATE:  
APPLICATION NUMBER: 07/833,431

FILING DATE: February 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: 96700/238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 286-0854 or 286-0082  
TELEFAX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 52297  
TYPE: nucleotide  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: phage genome sequence  
DESCRIPTION: phage genome sequence  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: not applicable.  
ORIGINAL SOURCE:  
ORGANISM: mycobacteriophage L5  
STRAIN: not applicable  
INDIVIDUAL ISOLATE: L5  
DEVELOPMENTAL STAGE: not applicable  
HAPLOTYPE: not applicable  
CELL TYPE: not applicable  
CELL LINE: not applicable  
ORGANELLE: not applicable  
IMMEDIATE SOURCE: mycobacteriophage L5 particles  
POSITION IN GENOME: entire genome  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS: Hatfull and Sarkis  
TITLE: DNA Sequence, Structure and Gene  
TITLE: Expression of Mycobacteriophage L5:  
TITLE: A Phage System for Mycobacterial  
TITLE: Genetics  
JOURNAL: Molecular Microbiology  
VOLUME: 7  
PAGES: 395-405  
DATE: 1993.

US-09-426-436-1

Query Match 3.2%; Score 34.6; DB 3; Length 52297;  
Best Local Similarity 49.2%; Pred. No. 4.7;  
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 431 TTGGCTAGACGACCAAGAACCCATTGACGCTTCGCGATCCAACTATCGGTGATC 490  
Db 18562 TTGCGCTGTTGAGGATGACGACCCCAAGACGATGTCGCTTCAGGGCTCGGTGGGAC 18503

Qy 491 AATTCGGACACCGCGGACTCGGACTTATCCAGGCTTCGTGTCGACCTTCG 550  
Db 18502 AAGTAGCGAGCGCGGGTAGTTCCTCCCGGCTCTTCAACACGACCTTGACGGAGGGTGC 18443

Qy 551 GATTACCTACGTCCTCACTTGGAGCCCTATCTCAGGACCGAGGAAGTTGCGCAACTGC 610  
Db 18442 TAGAAACACCTCCAGCTGTGTGGCCAGGTACACACCTTGTCACCGGTGCTCAAGTTG 18383

Qy 611 AAGCG 615  
Db 18382 AAACG 18378

RESULT 14  
US-08-705-557-1/c  
; Sequence 1, Application US/08705557

Patent No. 530061  
GENERAL INFORMATION:  
APPLICANT: William R. Jacobs, Jr.  
APPLICANT: Barry R. Bloom  
APPLICANT: Graham F. Hatfull  
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC  
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amster, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,557  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/057,531  
FILING DATE:  
APPLICATION NUMBER: 07/833,431  
FILING DATE: February 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: 96700/238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 52297  
TYPE: nucleotide  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: phage genome sequence  
DESCRIPTION: phage genome sequence  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: not applicable.  
ORIGINAL SOURCE:  
ORGANISM: mycobacteriophage L5  
STRAIN: not applicable  
INDIVIDUAL ISOLATE: L5  
DEVELOPMENTAL STAGE: not applicable  
HAPLOTYPE: not applicable  
CELL TYPE: not applicable  
CELL LINE: not applicable  
ORGANELLE: not applicable  
IMMEDIATE SOURCE: mycobacteriophage L5 particles  
POSITION IN GENOME: entire genome  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS: Hatfull and Sarkis  
TITLE: DNA Sequence, Structure and Gene  
TITLE: Expression of Mycobacteriophage L5:  
TITLE: A Phage System for Mycobacterial  
TITLE: Genetics  
JOURNAL: Molecular Microbiology  
VOLUME: 7  
PAGES: 395-405  
DATE: 1993.



PAGES: 395-405  
DATE: 1993  
US-08-705-557-1

Query Match  
Best Local Similarity 3.2%; Score 34.6; DB 4; Length 52297;  
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
QY 431 TTCGGCTAGACACACCAAGAACCCATGACGTCTTCGGATCCAACTATCGGTGTATC 490  
DB |||||  
18562 TTCGGCTGCTTGGAGTACGACACCCAAAGACGATGTCTCGTTCAGGGCTCGGTGGGAC 18503  
QY 491 AATTCGCGACACCGCGGACTCGCGACTCTATTCAGGCTTCGTGTCGACCTTCG 550  
DB |||||  
18502 AAGTAGCGAGCGCGGGTAGTTCGCCGCTTTCACACGACCTTGACGGGAGGTCG 18443  
QY 551 GATTACCTACCGTCCAACTTGAGGCTATCTCAGGACCGAGAAAGTTGGCGCACTGGC 610  
DB |||||  
18442 TAGAACACACCTCCAGCTCTGTGCGGAGGTACACACCTGTACCGGTGCTCAAGTTG 18383  
QY 611 AAGCG 615  
DB |||||  
18382 AAGCG 18378

RESULT 15  
US-09-198-452A-1/C  
Sequence 1, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Griffais, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
TITLE OF INVENTION: and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 1  
LENGTH: 1230025  
TYPE: DNA  
ORGANISM: Chlamydia pneumoniae  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(15000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (15001)..(30000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (30001)..(45000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (45001)..(60000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (60001)..(75000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (75001)..(90000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (90001)..(105000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (105001)..(120000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (120001)..(135000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (135001)..(150000)  
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc feature  
LOCATION: (150001)..(165000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (165001)..(180000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (180001)..(195000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (195001)..(210000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (210001)..(225000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (225001)..(240000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (240001)..(255000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (255001)..(270000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (270001)..(285000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (285001)..(300000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (300001)..(315000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (315001)..(330000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (330001)..(345000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (345001)..(360000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (360001)..(375000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (375001)..(390000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (390001)..(405000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (405001)..(420000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (420001)..(435000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (435001)..(450000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (450001)..(465000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (465001)..(480000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (480001)..(495000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (495001)..(510000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature

LOCATION: (510001)..(525000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (525001)..(540000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (540001)..(555000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (555001)..(570000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (570001)..(585000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (585001)..(600000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (600001)..(615000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (615001)..(630000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (630001)..(645000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (645001)..(660000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (660001)..(675000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (675001)..(690000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (690001)..(705000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (705001)..(720000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (720001)..(735000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (735001)..(750000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (750001)..(765000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (765001)..(780000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (780001)..(795000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (795001)..(810000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (810001)..(825000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (825001)..(840000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (840001)..(855000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (855001)..(870000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (870001)..(885000)

; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (885001)..(900000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature  
LOCATION: (900001)..(915000)  
; OTHER INFORMATION: n=a or c or g or t  
; NAME/KEY: misc feature

Query Match 3.1%; Score 33.2; DB 4; Length 1230025;  
Best Local Similarity 56.4%; Pred. No. 69;  
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
Qy 929 ACGACGAGAACGTTCCCAAACTAGAGGGAATCGGAAAATTGGACCCCATGGTTGGCT 988  
Db 751086 ATGACATAGTAATATCCGATATCAAAAGTTATCCAGACATTTGGTCATTTATTAGCAC 751027  
Qy 989 GCGAAATAAACAAAGGCGCTGACCCAGGATAGTGAACCTTGTGGAACCTGGTT 1038  
Db 751026 GCCATTACGCAAGTCAGAGAGATATGTTTTTGTATATTCAGACAGTGGCT 750977  
Search completed: March 8, 2004, 01:21:58  
Job time : 138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 00:23:29 ; Search time 468 seconds  
(without alignments)  
8487.118 Million cell updates/sec

Title: US-09-993-777-1\_COPY\_9874\_10959

Perfect score: 1086

Sequence: 1 ATGGCATCCTACTTGGAC.....CCCCGACTCAATAAATG 1086

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828715029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	13473	10	US-09-994-064-1
2	1084.4	99.9	3605	9	US-09-881-457A-3
3	1084.4	99.9	18913	10	US-09-994-064-59
4	343.3	31.6	534	10	US-09-994-064-18
5	38.2	3.5	777	12	US-10-282-122A-26478
6	38.2	3.5	780	12	US-10-282-122A-28211
7	37.6	3.5	6310	12	US-10-221-613-230
8	36.4	3.4	9025608	14	US-10-156-761-1
9	36	3.3	1049	14	US-10-123-155-358
10	36	3.3	1049	14	US-10-146-731-358
11	36	3.3	1049	14	US-10-140-472-358
12	36	3.3	1049	14	US-10-141-761-358
13	36	3.3	1049	14	US-10-142-885-358
14	36	3.3	1049	14	US-10-158-790-358
15	36	3.3	1049	15	US-10-137-871-358

C 16	36	3.3	1049	15	US-10-140-923-358	Sequence 358, App
C 17	36	3.3	1049	15	US-10-141-756-358	Sequence 358, App
C 18	36	3.3	1049	15	US-10-141-759-358	Sequence 358, App
C 19	36	3.3	1049	15	US-10-140-805-358	Sequence 358, App
C 20	36	3.3	1049	15	US-10-140-864-358	Sequence 358, App
C 21	35.6	3.3	1350	14	US-10-156-761-6893	Sequence 6893, App
C 22	35.2	3.2	19718	12	US-10-158-844-99	Sequence 99, Appl
C 23	35	3.2	653	14	US-10-184-644-402	Sequence 402, App
C 24	35	3.2	653	14	US-10-184-634-402	Sequence 402, App
C 25	35	3.2	1263	15	US-10-369-493-47306	Sequence 47306, A
C 26	35	3.2	1284	14	US-10-156-761-6164	Sequence 6164, Ap
C 27	35	3.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
C 28	34.6	3.2	475	13	US-10-040-739-899	Sequence 899, App
C 29	34.6	3.2	493	10	US-09-918-995-21276	Sequence 21276, A
C 30	34.6	3.2	1620	15	US-10-094-749-1549	Sequence 1549, Ap
C 31	34.4	3.2	1563	14	US-10-156-761-3340	Sequence 3340, Ap
C 32	34.2	3.1	802	14	US-10-184-644-312	Sequence 312, App
C 33	34.2	3.1	802	14	US-10-184-634-312	Sequence 312, App
C 34	34	3.1	2855	9	US-09-764-970-594	Sequence 594, App
C 35	34	3.1	2855	14	US-10-125-540-594	Sequence 594, App
C 36	33.8	3.1	500	14	US-10-063-685-84	Sequence 84, Appl
C 37	33.4	3.1	1581	15	US-10-369-493-40657	Sequence 40657, A
C 38	33.2	3.1	498	11	US-09-864-408A-3455	Sequence 3455, Ap
C 39	33.2	3.1	777	14	US-10-022-832-1	Sequence 1, Appli
C 40	33.2	3.1	777	15	US-10-312-273-56	Sequence 56, Appl
C 41	33.2	3.1	1296	15	US-10-260-238-92	Sequence 92, Appl
C 42	33.2	3.1	1442	12	US-10-425-114-27719	Sequence 27719, A
C 43	33.2	3.1	1230025	15	US-10-289-762-1	Sequence 1, Appli
C 44	32.8	3.0	1215	14	US-10-156-761-6851	Sequence 6851, Ap
C 45	32.6	3.0	2010	14	US-10-431-304-3	Sequence 3, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-994-064-1

; Sequence 1, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: WILG, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/469,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 391-0525

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

LENGTH: 13473 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: DNA (genomic)	
HYPOTHETICAL: NO	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 1059..2489	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 2575..4107	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 4113..4445	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 4609..5487	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 5697..8654	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 9874..10962	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 11159..12658	
FEATURE:	
NAME/KEY: CDS	
LOCATION: 12665..13447	
US-09-994-064-1	
Query Match	
Best Local Similarity 100.0%; Score 1086; DB 10; Length 13473;	
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1	ATGGCATCGTACTTGGAACTCTGCTCTCTCTGCGCGAGCTCGACACCTTCGGCGCG 60
DB 9874	ATGGCATCGTACTTGGAACTCTGCTCTCTCTGCGCGAGCTCGACACCTTCGGCGCG 9933
QY 61	ATGGGAATCGTATCACTGGAATCAGTCTCCGCCAGGATTGACGACATCATCGTG 120
DB 9934	ATGGGAATCGTATCACTGGAATCAGTCTCCGCCAGGATTGACGACATCATCGTG 9993
QY 121	ATCGTCGCGCTCGCGCCGAGCTACAATTCAGTCACTGAGCTATTTTTCGCTGGCCAG 180
DB 9994	ATCGTCGCGCTCGCGCCGAGCTACAATTCAGTCACTGAGCTATTTTTCGCTGGCCAG 10053
QY 181	AGACCCCAAAACCTACTCAGGAACCGTCGCGCTCGCGTTTCGCTCTGATATACAAAC 240
DB 10054	AGACCCCAAAACCTACTCAGGAACCGTCGCGCTCGCGTTTCGCTCTGATATACAAAC 10113
QY 241	CAGTCTACAGGAACCTTAGCGAGAGCGCTTTGAAATTCGACTCATCGATCGCTTCT 300
DB 10114	CAGTCTACAGGAACCTTAGCGAGAGCGCTTTGAAATTCGACTCATCGATCGCTTCT 10173
QY 301	GTCTTTGCGGTGTAAGTACCGAGTACAGTCTCCGCTCGAAGAGACTAACCGGA 360
DB 10174	GTCTTTGCGGTGTAAGTACCGAGTACAGTCTCCGCTCGAAGAGACTAACCGGA 10233
QY 361	CCTCCACACCGCTTAAAGTCACTATAGAAATCTCGTCCGACGAGCGGATGTC 420
DB 10234	CCTCCACACCGCTTAAAGTCACTATAGAAATCTCGTCCGACGAGCGGATGTC 10293
QY 421	TACGTAATTTGCGGTAGACACACCAAGAACCCATTGACGCTTCGCGATCCAATA 480
DB 10294	TACGTAATTTGCGGTAGACACACCAAGAACCCATTGACGCTTCGCGATCCAATA 10353
QY 481	TCGGTGATCAATTCGGGAACACCGCGGACTCGCGACTCTATTCCAGGCTTCGTGT 540
DB 10354	TCGGTGATCAATTCGGGAACACCGCGGACTCGCGACTCTATTCCAGGCTTCGTGT 10413
QY 541	CGACCTTCGGATTACCTACCGTCCAACTTGAGGCTATCTCAGGACCGAGGAAGTTGG 600

RESULT 2

US-09-881-457A-3  
Sequence 3, Application US/09881457A  
Patent No. US20020081316A1  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D  
APPLICANT: Cook, Stephanie M  
APPLICANT: Wild, Martha A  
TITLE OF INVENTION: No. US20020081316A1el Avian Herpes Virus and Uses Thereof  
FILE REFERENCE: SY01105K10KQK  
CURRENT APPLICATION NUMBER: US/09/881,457A  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: 09/426,352  
PRIOR FILING DATE: 1999-10-25  
PRIOR APPLICATION NUMBER: 08/804,372  
PRIOR FILING DATE: 1997-02-21  
PRIOR APPLICATION NUMBER: PCT/US95/10245  
PRIOR FILING DATE: 1995-08-09  
PRIOR APPLICATION NUMBER: 08/663,566  
PRIOR FILING DATE: 1996-06-13  
PRIOR APPLICATION NUMBER: 08/288,065  
PRIOR FILING DATE: 1994-08-09  
PRIOR APPLICATION NUMBER: PCT/US93/05681  
PRIOR FILING DATE: 1993-06-14  
PRIOR APPLICATION NUMBER: 08/023,610  
PRIOR FILING DATE: 1993-02-26  
PRIOR APPLICATION NUMBER: 07/898,087  
PRIOR FILING DATE: 1992-06-12  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn ver. 2.1  
SEQ ID NO 3  
LENGTH: 3605  
TYPE: DNA

```
; ORGANISM: Infectious Laryngotracheitis Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (585)..(1889)
; OTHER INFORMATION: ILTV glycoprotein D
; NAME/KEY: CDS
; LOCATION: (1997)..(3085)
; OTHER INFORMATION: ILTV glycoprotein I
US-09-881-457A-3

Query Match          99.9%; Score 1084.4; DB 9; Length 3605;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCGATCGCTACTTGGAACTCTGGCTCTCCCTTGGCGGACGCTGCGACCCCTTGGCGCG 60
Db 1997 ATGCGATCGCTACTTGGAACTCTGGCTCTCCCTTGGCGGACGCTGCGACCCCTTGGCGCG 2056

QY 61 ATGGGATCGTGATCACTGGAATCACTGCTCCGCCAGGATTGACGACGATCACATCGTG 120
Db 2057 ATGGGAATCGTGATCACTGGAATCACTGCTCCGCCAGGATTGACGACGATCACATCGTG 2116

QY 121 ATCGTCGCGCTCGCCCGGAGCTACAATTCAACTCAGCTATTTTTCATGCTGCGCCAG 180
Db 2117 ATCGTCGCGCTCGCCCGGAGCTACAATTCAACTCAGCTATTTTTCATGCTGCGCCAG 2176

QY 181 AGACCCCAAAACCCCTACTCAGAAACGCTCGCGTTCGGTCTCGATATAACAAC 240
Db 2177 AGACCCCAAAACCCCTACTCAGAAACGCTCGCGTTCGGTCTCGATATAACAAC 2236

QY 241 CAGTGCTACAGAACTTAGCGAGAGCGTTTGAATAATGCACTCATGCTGCTTCT 300
Db 2237 CAGTGCTACAGAACTTAGCGAGAGCGTTTGAATAATGCACTCATGCTGCTTCT 2296

QY 301 GTTTTGTGCGGTGTAAGTACCGGAGTACAGTCTCGCGCTCGAACAGACTAACCGGA 360
Db 2297 GTTTTGTGCGGTGTAAGTACCGGAGTACAGTCTCGCGCTCGAACAGACTAACCGGA 2356

QY 361 CTCCACACCGCTTAAAGTCACTATACGAATCTCTCGCGGACGACGCGGATGTC 420
Db 2357 CTCCACACCGCTTAAAGTCACTATACGAATCTCTCGCGGACGACGCGGATGTC 2416

QY 421 TAGCTAATTGTTGCGGTAGACGACACCAAGAACCCATTGAGTCTTCCGATCCAACTA 480
Db 2417 TAGCTAATTGTTGCGGTAGACGACACCAAGAACCCATTGAGTCTTCCGATCCAACTA 2476

QY 481 TCGGTGTATCAATTCGGAACACCGCGGACTCTGCGGACTCTATTCGAAGGCTTCGT 540
Db 2477 TCGGTGTATCAATTCGGAACACCGCGGACTCTGCGGACTCTATTCGAAGGCTTCGT 2536

QY 541 CGCACCTTCGGATTACCTACCGTCCAACTTGAGGCTATCTCAGGACCGAGGAAGTTGG 600
Db 2537 CGCACCTTCGGATTACCTACCGTCCAACTTGAGGCTATCTCAGGACCGAGGAAGTTGG 2596

QY 601 CGCAATGCGCAAGCTACGTTGCGACGAGGCGACGACGACGAGCGCGAGCGCAAC 660
Db 2597 CGCAATGCGCAAGCTACGTTGCGACGAGGCGACGACGACGAGCGCGAGCGCAAC 2656

QY 661 CCGACGCGCTCACTGCAACAGCGCTCCGAACCTTGAAGCGGAACACTTTACCTTTCC 720
Db 2657 CCGACGCGCTCACTGCAACAGCGCTCCGAACCTTGAAGCGGAACACTTTACCTTTCC 2716

QY 721 TGCTAGAAATGGCGTGAATCATTAAGAACCGACACCGCGAAAACGAAATTCAAACGTT 780
Db 2717 TGCTAGAAATGGCGTGAATCATTAAGAACCGACACCGCGAAAACGAAATTCAAACGTT 2776

QY 781 ACTGTCGCTCGGGACAAATGAGCCCTACGTAATTTGGGTAACTGCGTGGCTCGCGT 840
Db 2777 ACTGTCGCTCGGGACAAATGAGCCCTACGTAATTTGGGTAACTGCGTGGCTCGCGT 2836

QY 841 AGCGCAACGATCGGCTCTGTCATTGTAATTTTCATCGTCAACGAAACATGTGCAACCG 900
Db 2837 AGCGCAACGATCGGCTCTGTCATTGTAATTTTCATCGTCAACGAAACATGTGCAACCG 2896
```

RESULT 3  
US-09-994-064-59  
; Sequence 59, Application US/09994064  
; Publication No. US20030082788A1  
; GENERAL INFORMATION:  
; APPLICANT: Wild, Martha A.  
; APPLICANT: Cochran, Mark D.  
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.25  
; CURRENT APPLICATION DATA: US/09/994,064  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 08/468,190  
; APPLICATION NUMBER: 08/468,190  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 39116-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18913 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: N  
; ANTI-SENSE: N  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 697..1533  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: complement (1900..2784)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: CDS

QY 901 CACCGAAATTTAGACACGGTCTCCCAAGACGACGAAACGTTCCCAAACTAGAGGAA 960  
Db 2897 CACCGAAATTTAGACACGGTCTCCCAAGACGACGAAACGTTCCCAAACTAGAGGAA 2956  
QY 961 TCGCGAAATTTGGACCCATGTTGCGTGCAGATAAACAAGGCGCTGACCGAGTAGT 1020  
Db 2957 TCGCGAAATTTGGACCCATGTTGCGTGCAGATAAACAAGGCGCTGACCGAGTAGT 3016  
QY 1021 GAATTTGTGGAACCTGGTTGCGATTGTTAACCCGCTCGCTGCTAGCTGCCCGACTCAATA 1080  
Db 3017 GAATTTGTGGAACCTGGTTGCGATTGTTAACCCGCTCGCTGCTAGCTGCCCGACTCAATA 3076  
QY 1081 AAAATG 1086  
Db 3077 AAAATG 3082

LOCATION: complement (2916...3605)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3694...5124  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 5210...7081  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7245...8123  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 8333...11290  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 11098...12402  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 12510...13598  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 13792...15291  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 15298...16080  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16129...17013  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: complement (17380...18216)  
OTHER INFORMATION:  
US-09-994-064-59

Query Match 99.9%; Score 1084.4; DB 10; Length 18913;  
Best local Similarity 99.9%; Pred. No. 0;  
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCATCGTACTTGGAACTCTGGCTCTCCCTTCCGCGAGCTCGCACCCCTTCGGGCGG 60  
DB 12511 ATGGCATCGTACTTGGAACTCTGGCTCTCCCTTCCGCGAGCTCGCACCCCTTCGGGCGG 12570

QY 61 ATGGGAATCGTGATCACTTGGAAATCACTCTCCGCGAGGATGACGACGATCACTCGTG 120  
DB 12571 ATGGGAATCGTGATCACTTGGAAATCACTCTCCGCGAGGATGACGACGATCACTCGTG 12630

QY 121 ATCGTCGGCTCGCCCGAGCTTACAATCACTGACGCTATTTTCATGCTGCTGGCCAG 180  
DB 12631 ATCGTCGGCTCGCCCGAGCTTACAATCACTGACGCTATTTTCATGCTGCTGGCCAG 12690

QY 181 AGACCCCAACCACTTACTCAGGACCGTCCGCTCGGTTTCGCTGATATACAAC 240  
DB 12691 AGACCCCAACCACTTACTCAGGACCGTCCGCTCGGTTTCGCTGATATACAAC 12750

QY 241 CAGTGTACCAAGAACTTAGCGAGGAGCGCTTGAATTTGCATCTCATCGCTCTCT 300  
DB 12751 CAGTGTACCAAGAACTTAGCGAGGAGCGCTTGAATTTGCATCTCATCGCTCTCT 12810

QY 301 GTTTTTTGGCTGTAAGTGAACGAGTACAGTTCTCCGCTCGAAGCACTAACCGGA 360  
DB 12811 GTTTTTTGGCTGTAAGTGAACGAGTACAGTTCTCCGCTCGAAGCACTAACCGGA 12870

QY 361 COTCCACACCCGTTTAAAGTCTACTATACGAAATCTCTCGTCCGACGACGCGGATGTT 420

Db 12871 CCTCCACACCGTTTAACTCACTATAGAAATCTCTGTCGGAACGACAGCGGATGTT 12930

QY 421 TACGTAATTTGTTGGCTAGACGACACAAAGAACCCATTGACGTTCTTCGCATCAACTA 480

Db 12931 TACGTAATTTGTTGGCTAGACGACACAAAGAACCCATTGACGTTCTTCGCATCAACTA 12990

QY 481 TCGGTGTATCAATTCGCGAACACCGCCGCGACTCGCGACTCTATTCGAGGCTTCGTGT 540

Db 12991 TCGGTGTATCAATTCGCGAACACCGCCGCGACTCGCGACTCTATTCGAGGCTTCGTGT 13050

QY 541 CGCACCTTCGGATTACCTTACCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAGTTGG 600

Db 13051 CGCACCTTCGGATTACCTTACCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAGTTGG 13110

QY 601 CGCAACTGGCAAGCGTACGTTGCCACGAGGCGACGACGACGCGCGAGCGCAACACC 660

Db 13111 CGCAACTGGCAAGCGTACGTTGCCACGAGGCGACGACGACGCGCGAGCGCAACACC 13170

QY 661 CGGACGCGCTCACTTGCAACGAGCGCTCCGAACTTGAAGCGGACACTTTTACCTTTCC 720

Db 13171 CGGACGCGCTCACTTGCAACGAGCGCTCCGAACTTGAAGCGGACACTTTTACCTTTCC 13230

QY 721 TGGCTAGAAATGGCGTGATCATTTACGAAACCGACCGCGAAACGAAATTCGAAACGTT 780

Db 13231 TGGCTAGAAATGGCGTGATCATTTACGAAACCGACCGCGAAACGAAATTCGAAACGTT 13290

QY 781 ACTGTCCGTCTCGGACAAATGAGCCCTACGCTTAATTTGGGGTAAACGCTGCTCCCTCG 840

Db 13291 ACTGTCCGTCTCGGACAAATGAGCCCTACGCTTAATTTGGGGTAAACGCTGCTCCCTCG 13350

QY 841 AGCGCAACGATCGGCTCGTCAATTTGTAATTTCCATCTGTCACGAAACATGTGCAACCGG 900

Db 13351 AGCGCAACGATCGGCTCGTCAATTTGTAATTTCCATCTGTCACGAAACATGTGCAACCGG 13410

QY 901 CACCGAAATTTAGACACGCTCTCGCAAGACGAGGAAAGCTTCCCAAACTAGAGGGGAA 960

Db 13411 CACCGAAATTTAGACACGCTCTCGCAAGACGAGGAAAGCTTCCCAAACTAGAGGGGAA 13470

QY 961 TCGCGAAATTTGGACCCCATGTTGCGTCCGAAATTAACGAGGCGCTGACACGATAGT 1020

Db 13471 TCGCGAAATTTGGACCCCATGTTGCGTCCGAAATTAACGAGGCGCTGACACGATAGT 13530

QY 1021 GAACCTTGGAACTGTTGCGATTTTAAACCGCTTCCGCTTAAGCTCGCCGACTCAATA 1080

Db 13531 GAACCTTGGAACTGTTGCGATTTTAAACCGCTTCCGCTTAAGCTCGCCGACTCAATA 13590

QY 1081 AAAATG 1086

Db 13591 AAAATG 13596

RESULT 4  
US-09-994-064-18/c  
; Sequence 18, Application US/09994064  
; Publication No. US20030082788A1  
; GENERAL INFORMATION:  
; APPLICANT: Wild, Martha A.  
; APPLICANT: Cochran, Mark D.  
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/994,064  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/469,190  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 391116-A  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 534 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..533  
US-09-994-064-18

Query Match 31.6%; Score 343; DB 10; Length 534;  
Best Local Similarity 100.0%; Pred. No. 4.1e-109;  
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAAAACGAAATTCACACGTTACTGTCGCGTCTCGGACACATGAG 803  
DB 534 TTACGACCGACACCGCGCAAAACGAAATTCACACGTTACTGTCGCGTCTCGGACACATGAG 475  
QY 804 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGGTGAGCGCAACCGATCGGCTCGTCAT 863  
DB 474 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGGTGAGCGCAACCGATCGGCTCGTCAT 415  
QY 864 TGTAATTTCCATCGTCACAGAAACATGTGCAACCGCGCAACCGAAATTTAGACCGGTCTC 923  
DB 414 TGTAATTTCCATCGTCACAGAAACATGTGCAACCGCGCAACCGAAATTTAGACCGGTCTC 355  
QY 924 GCAAGACGACGACGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGCT 983  
DB 354 GCAAGACGACGACGACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGCT 295  
QY 984 TCGGTCGGAATAAACAAGCGCGCTGACCGAGGATAGTGAATTTGGAACTGTTGGCAT 1043  
DB 294 TCGGTCGGAATAAACAAGCGCGCTGACCGAGGATAGTGAATTTGGAACTGTTGGCAT 235  
QY 1044 TGTTAACCCGCTGCGCTAAGCTCGCGCGCACTCAATAAAATG 1086  
DB 234 TGTTAACCCGCTGCGCTAAGCTCGCGCGCACTCAATAAAATG 192

RESULT 5  
US-10-282-122A-26478/c  
Sequence 26478, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 26478  
LENGTH: 777  
TYPE: DNA  
ORGANISM: Mycobacterium bovis  
US-10-282-122A-26478

Query Match 3.5%; Score 38.2; DB 12; Length 777;  
Best Local Similarity 58.3%; Pred. No. 0.036;  
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCTTCGCGCGATGGGAATCGTGATCACTGGAATCAACG 88  
DB 544 TCCACGCGCCACCTCGCACCTTCGCGCGATGGGAATCGTGATCACTGGAATCAACG 485  
QY 89 TCTCGCGCAGATTGACGACGATCACATCGTGCGGCGCTCGCGCGCAAGC 143  
DB 484 TCGGCGCGCAGCAGGTCCTCAGCGCGGACCGCGCGCAAGCAGCGCGC 430

RESULT 6  
US-10-282-122A-28211/c  
Sequence 28211, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23

Query Match 3.5%; Score 37.6; DB 12; Length 6310;  
Best Local Similarity 49.0%; Pred. No. 0.18;

RESULT 9  
US-10-123-155-358/c



```
; Sequence 358, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Demoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-358

Query Match      3.3%; Score 36; DB 14; Length 1049;
Best Local Similarity 4.5%; Pred. No. 0.25;
Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

QY 270 CTTTGAATTCGACTCATCGATCGTCTCTCTGTTTGTGCGGTGTAAGTGCACGAGTA 329
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
818 C.CBGBAHT.R..AB..ADNBWB.WBTHGAG..YND..KM.NN.BN...SAM..K.. 759
QY 330 CACCTTCCTCGCTCGACGAGTAACCGGACCTCCACACCGCTTAAAGTCACTATACG 389
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
758 MNSS.H.RY...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
QY 390 AAATCCTC-GTCCGAAACGACAGAGTGTCTACGTAATGTTCGGCTAGACGACCA 448
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
698 G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H...M..NM....RYNH 639
QY 449 AAGACCCATTGACGCTCTTCGCGATCCCACTATCGGTGTATCAATTCGGAACACGCG 508
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638 C.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHKK.M..BM.NMA.N.KDA.C.S..R 579
QY 509 CGACTCGGAGTCTATTCCAAGCTTCGTGTCGACCTTCGGATTACCTACGTCAC 568
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
578 DSNS.H.B..MD....TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM. 519
QY 569 TTGAGGCTATCTCAGGACGAGGAAAGTTGGCGCACTGGCAAGCTAGTTGCCACCG 628
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 S.D...HSSMB...SNMS.H.A.CRMGS.NBSK.ST.MNM.VGSTRMRY.RD....B 459
QY 629 AGGCCAGGACGACGAGCGGACACACCCCGCTGCTCACTGCAACACGAGCGCT 688
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 ...RS.BSAYNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B.. 399
QY 689 CCGAATTGAGCGGAACATTTACCTTTCCCTGGGTAGAAATGGCGTGATCATACG 748
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB....N.S.H....S.. 339
QY 749 AACGACACCGGAAACGAATTCACGTTACTGTCCGCTCGGACATGAGCCTA 808
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 D..MTHC..MT..N.S.H....MN.NM.WY..BD..SNSD.Y.BM..A.TH.TNB.... 279
```

```
QY 809 CCTTAATTTGGGTAAACCGTGGTCCGTCGTGAGGGGCAACGATCGGCTCTCATTTAA 868
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 .SNNMG.TG...TNRGY.GNCS.H.....NN.HH...MT.KNNR.R..A.AS..BA.B 219
QY 869 TTTCCATCGTCACAGAAACATGTGCACCCCGCGAAATTTAGACACGCGTCTCGCAAG 928
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 TABNNHM.S.BM.MA.N..THM..SRBRG.NYRGN.C.R....NT..A.N.MY.S..N 159
QY 929 ACGACGAAAGAGTTCCTCCAAACTAGAGGGAATCGGAAATTTGGACCCATGTTGCGT 988
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 NT..S....S...C.....NCH.R.SM.RA.CS.SY.M...YM.GKNNMSC....BGN 99
QY 989 GCGAAATTAACAAGGCGCTGACCGAGGATAGTGAACCTGTGGAATGTTGCGATTGTA 1048
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDNM.BH.A 39

RESULT 10
US-10-146-731-358/c
; Sequence 358, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Demoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-358

Query Match      3.3%; Score 36; DB 14; Length 1049;
Best Local Similarity 4.5%; Pred. No. 0.25;
Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

QY 270 CTTTGAATTCGACTCATCGATCGTCTCTCTGTTTGTGCGGTGTAAGTGCACGAGTA 329
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
818 C.CBGBAHT.R..AB..ADNBWB.WBTHGAG..YND..KM.NN.BN...SAM..K.. 759
QY 330 CACGTTCTCCGCTCGAAGACTAACCGGACCTCCACACCGCTTAAAGTCACTATACG 389
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
758 MNSS.H.RY...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
QY 390 AAATCCTC-GTCCGAAACGACGAGGATGTCTACGTAATGTTCGGCTAGACGACCA 448
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
698 G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H...M..NM....RYNH 639
QY 449 AAGACCCATTGACGCTCTTCGCGATCCCACTATCGGTGTATCAATTCGGAACACGCG 508
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638 C.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHKK.M..BM.NMA.N.KDA.C.S..R 579
QY 509 CGACTCGGAGTCTATTCCAAGCTTCGTGTCGACCTTCGGATTACCTACGTCAC 568
```

Db.	578	DSNSS.H.B..MD.....TASD..H.VNNS.H.RY..T.....SCN.A.S..NCS.N.GM..519
Qy	569	TTGAGCGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGCGCAAGCGTACGTTGGCCACGG 628
Db	518	.S.D..HSSMB...SNMS.H.A.CRMGRS.NBSK.ST.MNM.VGSTRMHRY.RD....B 459
Qy	629	AGGCCACGACGACCGAGCGCGAGCGCAACCCGAGCGCCGCTACTGCAACACAGCGCCT 688
Db	458	...RS.BSAVTNSG.CB.SSHCS.S.MVBS.H.BM.YM..M.KS.NT.M..NAC.H.B..399
Qy	689	CGGAACCTGAAGCGGAACACTTTACCTTTCCTCGGTAGAAAATGGCGTGATCAITACG 748
Db	398	N..ND..S.N.SM..M.BRCY.Y.M.SM.SS.T.S.NKSTYRB....N.S.H....S...339
Qy	749	AACGACACCGCGCAACGAAATTCAAACGTTACTGTCGCTCGGACAAATGAGCCCTA 808
Db	338	D..MTHC..MT..N.S.H....NM.NM.WY..BD..SNSD.Y.BM..A.TH.TNB....279
Qy	809	CGCTAATTTGGGTAAACGCTGGCTCCGCTCGTAGCGCAACGATCGGCGCTCGTCATTGTAA 868
Db	278	.SNNMG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNMR.R..A.AS..BA.B 219
Qy	869	TTTCCATCGTCACCGAGAAACATGTGCACCCCGCAGCGCAAAATTAGACACGCTCTCGAAG 928
Db	218	TABNNHM.S.BM.NA.N..THM..SRBGR.NYRRGN.C.R....NT..A.N.MY.S..N 159
Qy	929	ACGACGAAGAACGTTCCCAAACTAGAAGGAAATCGCGAAAATTTGGACCCATGTTGCGT 988
Db	158	NT..S....S..C.....NCH.R.SM.RA.CS.SY.M..YM.GKNMSC....BGN 99
Qy	989	CGCAAAATAACAAGGGCGCTACACAGGATAGTGAACCTTTGTGGAACCTCGTTGCGATGTTA 1048
Db	98	GY.H..B.DH.YD.ST.S.H...DN.A.A.NAANA..CC...A.DMAGHB.BNNM.BH.A 39

## RESULT 11

US-10-140-472-358/c  
; Sequence 358, Application US/10140472  
; Publication No. US20030138888A1

GENERAL INFORMATION:

APPLICANT:	Baker, Kevin P.
APPLICANT:	Beresini, Maureen
APPLICANT:	DeForge, Laura
APPLICANT:	Desnoyers, Luc
APPLICANT:	Filvaroff, Ellen
APPLICANT:	Gao, Wei-Qiang
APPLICANT:	Gerritsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Gurney, Austin L.
APPLICANT:	Sherwood, Steven
APPLICANT:	Smith, Victoria
APPLICANT:	Stewart, Timothy A.
APPLICANT:	Tumas, Daniel
APPLICANT:	Watanabe, Colin K
APPLICANT:	Wood, William
APPLICANT:	Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME	

Query Match 3.3%; Score 36; DB 14; Length 1049;  
Best Local Similarity 4.5%; Pred. No. 0.25;

Matches	35; Conservative	215; Mismatches	529; Indels	1; Gaps
QY	270	CTTTGAAATTCGACTCATCGATCGCTCTCTCTCTTTTTCGGCTCTAAAGTCACCGAGTA	329	
Db	818	C.CBGA BHAT.R..AB.ADNBWB.WBTHGAG..YND.D...KM.NN.BN...SAM..K..	759	
QY	330	CAGGTTCTCGGCTCGAAGACTAACCGGACTCCACACCGTTTAAAGCTCACTATACG	389	
Db	758	MNSS.H.RY...TH...RMA.SY..NNM...NM.SYGN.S.Y..BAA..NDS.H.A..NM	699	
QY	390	AAATCTCTC-GTCCGAACGACGCGGATGTTCTACGTAATTTCTCGGCTAGACGACCA	448	
Db	698	G..G..MMWS.SV.CNWT.S.NM.N..KCH.BCS...S.SNWS.H...M..NM...RYNH	639	
QY	449	AAGAACCATTGACGCTTTCGGCATCCCACTATCGGTGATCAATTCGCGAAGACGCGCG	508	
Db	638	C.YW.BH.DNCV...AY.S.S.KAYSASS.HHNKK.M..BM.NMA.N.KDA.C.S..R	579	
QY	509	CGACTCGGAGCTCTATCCCAAGGCTTCGTGTCGACCTTCGATTAACCTACGTCAC	568	
Db	578	DSNSS.H.B..MD...TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM.	519	
QY	569	TTGAGCGCTTCTCAGGACCGGAGGAAGTTGGCGCAACTGGCAAGCTGATGCCACG	628	
Db	518	S.D..HSNB...SNMS.H.A.CMRGS.NESK.ST.NMM.YGSYTRHRY.RD....B	459	
QY	629	AGGCCACGACGACGCGCGGAGGCGAACAACCCGACGCCCTCACTGCAACAGCGCCT	688	
Db	458	..RS..BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B..	399	
QY	689	CGAATCTGAAGGGAACACTTACTTTCCTCTGGCTAGAAAATGCGTGGATCATACG	748	
Db	398	N..ND..S.N.SM..M.BCY.Y..M.SM.SS.T.S.NKSTYRB...N.S.H...S...	339	
QY	749	AACCGACACCCGAAACGAAATTCAAAGCTTACTGCTCGTCTCGGCAAAATGAGCCCTA	808	
Db	338	D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.EM..A.TH.TNB....	279	
QY	809	CGCTAATTCGGGTAAACCGTGGTGGCTGCGTGGAGCGCAACGATCGGCCTCGTCAATGTA	868	
Db	278	SNNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R..A.AS..BA.B	219	
QY	869	TTTCCATCGTCAACGAAACATGTGCACCCCGACCGAAATTAGACACGGTCTCGAAG	928	
Db	218	TABNNHM.S.BM.MA.N..THM..SRBGR.NYRRGN.C.R...NT..A.N.MY.S..N	159	
QY	929	ACGACGAAGACGTTCCCAACTAGAGGGAATCGGAAAATTTGGACCCATGTTGCGT	988	
Db	158	NT..S...S...C.....NCH.R.SM.RA..CS.SY.M...YM.GKNMSC....BGN	99	
QY	989	GCGAAATAACAGGCGGCTGACCGAGATGTAACCTTGGAACCTGGTTCGATTTGA	1048	
Db	98	GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDNM.BH.A	39	

RESULT 12  
US-10-141-761-358/c  
; Sequence 358, Application US/10141761  
; Publication No. US20030148432A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.

## RESULT 12

US-10-141-761-358/C

Sequence 358, Application US/10141761

Publication No. US20030148432A1

GENERAL INFORMATION:

APPLICANT:	Baker, Kevin P.
APPLICANT:	Bereain, Maureen
APPLICANT:	DeForest, Laura
APPLICANT:	Deenoyers, Luc
APPLICANT:	Filvaroff, Ellen
APPLICANT:	Gao, Wei-Qiang
APPLICANT:	Gerritsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Gurney, Austin L.
APPLICANT:	Sherwood, Steven
APPLICANT:	Smith, Victoria
APPLICANT:	Stewart, Timothy A.

```
/ APPLICANT: Tumas,Daniel
/ APPLICANT: Watanabe,Colin K
/ APPLICANT: Wood,William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C198
/ CURRENT FILING DATE: 2002-05-08
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 358
/ LENGTH: 1049
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-141-761-358

Query Match      3.3%; Score 36; DB 14; Length 1049;
Best Local Similarity 4.5%; Pred. No. 0.25;
Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

QY 270 CTTTGAATAATGCACTATCATCGATCGTCTCTGTTTGTGGCTGTAAAGTACCGAGTA 329
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
818 C.CBGAHAT.R..AB.ADNBWB.WBTHGAG..YND...KM.NN.EN...SAM..K.. 759

QY 330 CACGTTCTCCGCTCGACACACTAACCGGACCTCCACCCGTTTAAGTCACTATACG 389
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
758 MNSS.H.R.Y...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699

QY 390 AAATCCTC-GTCCGAACGACAGCGGATGTTCTAGTAATGTTCCGCTAGACGACACCA 448
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
698 .G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S.SNWS.H...M..NM....RYNH 639

QY 449 AAGAACCAATGACGCTTCGCGATCCCACTATCGTGTATCAATTCGGAACACCGCG 508
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638 C.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHNKK.M..BM.NMA.N.KDA.C.S..R 579

QY 509 CGACTCGGGACTATTCCAAAGCTTCTGTCGACCTTCCGCTGATCACTATACGTCG 568
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
578 DNSS.H.B..MD....TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM. 519

QY 569 TTGAGCGCTTCTCAGGACCGGAGAAAGTTGGCGCACTGGCAAGCTAGCTGCGACCG 628
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 .S.D..HSSMB...SNMS.H.A.CMRGS.NBSK.ST.MNM.YGSTRMHR.YD....B 459

QY 629 AGGCCACGACGACCGCGGAGGACACACCCGACGCGCTACTGCAACGCGCT 688
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 ...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.N..NAC.H.B.. 399

QY 689 CCGAACTTGAACGGAACACTTTTACTTTCCTCGCTAGAAATGCGGTGATCATACG 748
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 N..ND..S.NM..M.BCY.Y..M.SM.SS.T.S.NKSTRB...N.S.H....S... 339

QY 749 AACCGACACCGCAACGAAATTCAAAGCTTACTGTCGCTCGGACCAATGAGCCCTA 808
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDMN.BH.A 39
```

## RESULT 13

US-10-142-885-358/c  
Sequence 358, Application US/10142885  
Publication No. US20030157604A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria A.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C248  
CURRENT APPLICATION NUMBER: US/10142,885  
Prior Filing Date: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-142-885-358

Query Match 3.3%; Score 36; DB 14; Length 1049;

Best Local Similarity 4.5%; Pred. No. 0.25; Mismatches 529; Indels 1; Gaps 1;

Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

```
QY 270 CTTTGAATAATGCACTATCATCGATCGTCTCTGTTTGTGGCTGTAAAGTACCGAGTA 329
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
818 C.CBGAHAT.R..AB.ADNBWB.WBTHGAG..YND...KM.NN.BN...SAM..K.. 759

QY 330 CACGTTCTCCGCTCGACACACTAACCGGACCTCCACCCGTTTAAGTCACTATACG 389
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
758 MNSS.H.RY...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699

QY 390 AAATCCTC-GTCCGAACGACAGCGGATGTTCTAGTAATGTTCCGCTAGACGACACCA 448
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
698 .G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S.SNWS.H...M..NM....RYNH 639

QY 449 AAGAACCAATGACGCTTCGCGATCCCACTATCGTGTATCAATTCGGAACACCGCG 508
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638 C.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHNKK.M..BM.NMA.N.KDA.C.S..R 579

QY 509 CGACTCGGGACTATTCCAAAGCTTCTGTCGACCTTCCGCTGATCACTATACGTCG 568
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
578 DNSS.H.B..MD....TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM. 519

QY 569 TTGAGCGCTTCTCAGGACCGGAGAAAGTTGGCGCACTGGCAAGCTAGCTGCGACCG 628
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 .S.D..HSSMB...SNMS.H.A.CMRGS.NBSK.ST.MNM.YGSTRMHR.YD....B 459

QY 629 AGGCCACGACGACCGCGGAGGACACACCCGACGCGCTACTGCAACGCGCT 688
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 ...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.N..NAC.H.B.. 399

QY 689 CCGAACTTGAACGGAACACTTTTACTTTCCTCGCTAGAAATGCGGTGATCATACG 748
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 N..ND..S.NM..M.BCY.Y..M.SM.SS.T.S.NKSTRB...N.S.H....S... 339

QY 749 AACCGACACCGCAACGAAATTCAAAGCTTACTGTCGCTCGGACCAATGAGCCCTA 808
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDMN.BH.A 39
```



Query Match	3.3%; Score 36; DB 15; Length 1049;
Best Local Similarity	4.5%; Pred. No. 0.25;
Matches	Conservative 215; Mismatches 529; Indels 1; Gaps 1;
Qy	270 CTTTGAATAATTGCACCTCATGCATCGTCTCTGTTTGTTCGGCTGTAAGTGACCGAGTA 329
Db	818 C.CEGBABT.R.AB.ADNBWB.WETHGAG..YND..KM.NN.BN..SAM..K..759
Qy	330 CAGCTTCTCGCCTCGAACAGACTAACCGGACCTCCACACCGGTTTAAGCTCACTATACG 389
Db	758 MNSS.H.RY..TH..RMA.SY..NNM..NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
Qy	390 AATCTCTC-GTCCGAACGACAGCGGGATGTTCTACGTAATTGTTGCGCTAGACGACCA 448
Db	698 G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S..SNMS.H...M..NM...RYNH 639
Qy	449 AAGAACCATTGACGCTCTTCGGATCCAACTATCGGTGTATCATTCGCGAAACACCGCG 508
Db	638 C.YW.BH.DNCY...AY.S.S.KAYSASS.HNENKK.M..BM.NMA.N.KDA.C.S..R 579
Qy	509 CGACTCGCGACTATTTCGAAGCCTTCGTGTCCGACCTTCGGATTACCTACGTCOAC 568
Db	578 DSNSS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S..NCS.N.GM. 519
Qy	569 TTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCACTGGCAAGCTGCTTGGCCACG 628
Db	518 S.D..HSSMB...SNMS.H.A.CRMGS.NBSK.ST.MNM.YGSIYRMHRY.RD...B 459
Qy	629 AGGCGACGACGACCGGCGGAGCGGCAACCCGAGCGCCGTCACATGCAACACCGCCT 688
Db	458 ...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B... 399
Qy	689 CCGAATTTGAACGGGAACATTTACCTTTCCCTGGCTAGAAAATGGCGGTGATCATTCG 748
Db	398 N..ND..S.N..SM..M.BRCY.Y..M.SM..SS.T.S.NKSTYRB...N.S.H...S... 339
Qy	749 AACCGACACCGCAACGAAATTCAAAGGTTACTGTCGCTCGGACCAATGAGCCCTA 808
Db	338 D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A.TH.TNB.... 279
Qy	809 CGCTAAATTGGGGTAACCGTGCTGCCGCTCGTGAACGCAACGATCGGCTCGTCAATTGAA 868
Db	278 ..SNNWG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNNR.R..A.AS..BA.B 219
Qy	869 TTTTCATGTCACAGAAACATGTGCACCCGACCGAAAATTAGACACGGTCTCGCAAG 928
Db	218 TABNNHM.S.BM.NA.N..THM..SRBGR.NYRGRN.C.R....NT..A.N.MY.S..N 159
Qy	929 ACGACGAAGAACGTTCCCAACTAGAAGGGAATTCGGAATAATTTGGACCCATGTTCCGT 988
Db	158 NT..S....S..C.....NCH.R.SM.RA.CS.SY.M..YM.GKNVSC.....BGN 99
Qy	989 GCGAAATAACAAGGGCGCTGACCGAGATGATGAACCTTGTGGAACTCGTTCGGATTGTTA 1048
Db	98 GY.H..B.DH.YD..ST.S.H..DN.A.A.NAANA..CC..A.DMHAGHB..BDNM..BH.A 39

Search completed: March 8, 2004, 02:39:41  
Job time : 487 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 22:42:19 ; Search time 3277 Seconds  
(without alignments)  
9896.347 Million cell updates/sec

Title: US-09-993-777-1\_COPY\_9874\_10959  
Perfect score: 1086  
Sequence: 1 ATGGCATCGTACTTGGAAC.....CGCCGACTCAATAAAATG 1086

Scoring table: IDENTITY NUC  
Gapop 10\_0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.2	4.0	493	10 BE777414	BE777414 MY-35-A-0
C 2	43.2	4.0	885	13 EX425603	EX425603 BX425603
C 3	42	3.9	681	9 AU076048	AU076048 AU076048
C 4	41.4	3.8	591	9 AU068783	AU068783 AU068783

5	41.4	3.8	765	28 AZ132997	AZ132997 OSJNB010
6	41.4	3.8	1053	29 CMS060TI	AL411100 T7 end of
7	41.2	3.8	453	14 CA704113	CA704113 wdkic.pk0
C 8	41	3.8	618	14 CB883239	CB883239 HQ01J24W
C 9	40.2	3.7	852	13 BX393687	BX393687 BX393687
C 10	39.8	3.7	1201	13 BX360624	BX360624 BX360624
C 11	39.4	3.6	328	29 CE179705	CE179705 tigr-gss-
C 12	39.4	3.6	551	29 CE554687	CE554687 tigr-gss-
C 13	38.6	3.6	577	14 CA255506	CA255506 SCEPFL418
C 14	38.6	3.6	717	14 CF877969	CF877969 trico077XP
15	38.6	3.6	718	14 CF865608	CF865608 trico002Xe
16	38.6	3.6	727	14 CF867578	CF867578 trico11X0
17	38.6	3.6	735	14 CF866505	CF866505 trico06X0
18	38.6	3.6	747	14 CF868489	CF868489 trico015X0
19	38.6	3.6	750	14 CF866359	CF866359 trico006X0
20	38.6	3.6	752	14 CF866952	CF866952 trico008Xn
21	38.6	3.6	772	14 CB906186	CB906186 trico077XP
22	38.6	3.6	775	14 CB895519	CB895519 trico002Xe
23	38.6	3.6	788	14 CB897619	CB897619 trico11X0
24	38.6	3.6	791	14 CB896481	CB896481 trico006X0
25	38.6	3.6	803	14 CB898569	CB898569 trico15X0
26	38.6	3.6	805	14 CB896951	CB896951 trico008Xn
27	38.6	3.6	807	14 CB896324	CB896324 trico006X0
28	38.6	3.6	1201	13 BX360624	BX360624 BX360624
C 29	38.2	3.5	575	12 BI352860	BI352860 GM21095.5
30	38.2	3.5	720	28 BZ896103	BZ896103 NARP6_012
31	38	3.5	732	14 CF868364	CF868364 trico14X1
32	38	3.5	788	14 CB898441	CB898441 trico14X1
33	37.8	3.5	865	14 CB899074	CB899074 trico17X0
34	37.8	3.5	865	14 CF868966	CF868966 trico17X0
C 35	37.8	3.5	925	29 CNS0091P	AL033013 prosophi1
36	37.6	3.5	693	14 CF882808	CF882808 trico087Xf
37	37.6	3.5	749	14 CB909128	CB909128 trico087Xf
38	37.6	3.5	1200	9 AL579901	AL579901 AL579901
39	37.4	3.4	1201	13 BX381961	BX381961 BX381961
C 40	37.2	3.4	591	14 CA181075	CA181075 SCAGST313
41	37	3.4	519	13 CA127859	CA127859 SCAGLR201
C 42	37	3.4	550	12 BM596053	BM596053 170006874
C 43	37	3.4	675	12 BM603983	BM603983 170006870
44	37	3.4	722	14 CF869283	CF869283 trico18Xm
C 45	37	3.4	731	12 BM585009	BM585009 170006872

#### ALIGNMENTS

RESULT 1  
BE777414/c  
LOCUS BE777414 493 bp mRNA linear EST 20-SEP-2000  
DEFINITION MY-35-A-04 PinfestansMY Phytophthora infestans cDNA, mRNA sequence.  
ACCESSION BE777414  
VERSION BE777414.1 GI:10231053  
KEYWORDS EST.  
SOURCE Phytophthora infestans (potato late blight agent)  
ORGANISM Phytophthora infestans  
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.

REFERENCE 1 (bases 1 to 493)

AUTHORS Kamoun, S., Hraber, P.T., Sobral, B.W.S., Nuss, D. and Govers, F.  
TITLE Initial assessment of gene diversity for the oomycete pathogen  
Phytophthora infestans based on expressed sequences

JOURNAL Fungal Genet. Biol. 28 (2), 94-106 (1999)

MEDLINE 20056376

PUBMED 10587472

COMMENT

Contact: Govers F

Laboratory of Phytopathology

Wageningen University

Binnenhaven 9, P.O.Box 8025, 6700 BE, Wageningen, The Netherlands

Tel: 31 317 483 138

Fax: 31 317 483 412

Email: Francine.Govers@medew.fyto.wau.nl.

Location/Qualifiers

1. .493

FEATURES source



/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="Z2735.62"  
/dev\_stage="flowering stage"  
/clone\_lib="Rice panicle at flowering stage"  
/note="Organ: panicle; Rice cDNA from panicle at flowering stage"

## ORIGIN

Query Match 3.9%; Score 42; DB 9; Length 681;  
Best Local Similarity 53.4%; Pred. No. 6.4;  
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
  
QY 560 CGTCCAACTTGGGCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCGAGCGTAGC 619  
DB 354 CCATCCGGGTTGAGAGTTCGTTTGAAGATGCTGACGTGGCGCAGGCGCTGCGCATGT 295  
  
QY 620 TTGCCAGGAGCCAGCAGCAGCGCGGAGCGGACACCGCCGCTCACTGCAA 679  
DB 294 CGGCAACCGAGCCACCGCGGAGCAGCGACGACGAGCGCGGCGGCGCTTACGCGCT 235  
  
QY 680 CCAGCGCCCTCCGAACCTTGAAGCGGAACTTACCTTTCCCTG 722  
DB 234 GCAGCGCCACCCACTTCGCGCTGCTCTGNACTTCGCTG 192

## RESULT 4

AU068783/c  
LOCUS AU068783 Rice callus Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone C50531\_102, mRNA sequence.

ACCESSION AU068783

VERSION AU068783.1 GI:5003634

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 591)

Yamamoto, K. and Sasaki, T.

Rice cDNA from callus (1998)

Unpublished (1998)

Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/

PROJECT = RGP.

PROECT = RGP.

Location/Qualifiers

FEATURES

source

1..591  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="RNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="C50531.102"  
/clone\_lib="Rice callus"  
/note="Vector: pBluescript II SK+; Site 1: SalI; Site 2:  
NotI; cDNA prepared from rice callus mRNAs by using  
oligo(dT) as a primer and ligating to the SalI-NotI site  
of pBluescript II SK+ phagemid."

## ORIGIN

Query Match 3.8%; Score 41.4; DB 9; Length 591;  
Best Local Similarity 53.4%; Pred. No. 8.7;  
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
  
QY 560 CGTCCAACTTGGGCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCGAGCGTAGC 619  
DB 260 CCATCCGGGTTGAGAGTTCGTTTGAAGATGCTGACGTGGCGCGCTGCGCATGT 201

QY 620 TTGCCAGGAGCCAGCAGCAGCGCGGAGCGCAACCCGAGCCCGTCACTGCAA 679  
DB 200 CGGCCACCGAGCCACCGCGCGGAGCAGCAGCAGCGCGGCGCTTGAAGCGCT 141  
  
QY 680 CCAGCGCTCCGAACCTTGAAGCGGAACTTACCTTTCCCTG 722  
DB 140 GCAGCGCCACCCACTTCGCGCTGCTCTGACCTTCGCGCTG 98

## RESULT 5

AZ132997

LOCUS AZ132997

DEFINITION OSJNB0108G22f CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone OSJNB0108G22f, genomic survey sequence.

ACCESSION AZ132997

VERSION AZ132997.1

KEYWORDS GSS.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 765)

Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Rice Genome

Unpublished (1998)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: GTAAACGACGCGCCAGTG

Class: BAC ends

High quality sequence start: 31

High quality sequence stop: 580.

Location/Qualifiers

FEATURES

source

1..765  
/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="genomic DNA"

/strain="Japonica"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNB0108G22f"

/tissue\_type="Leaf"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice BAC Library (EcoRI)"

/note="Vector: pBACindigo; Site 1: EcoRI; Site 2: EcoRI;

Rice is the most important food crop in the world. Half of

the world population, especially those inhabiting highly

populated areas of the humid tropics and subtropics, rely

on rice as their primary source of carbohydrate.

Monocotyledonous rice is a diploid plant (2n=24) with a

haploid genome equivalent of 431 Mbp (Arumuganathan and

Earle, 1991). The relatively small genome of rice, three

times larger than that of Arabidopsis, makes it suitable

for genomic studies. In order to facilitate positional

cloning, physical mapping and genome sequencing of rice,

we have constructed a BAC library from Oryza sativa,

Nipponbare variety using EcoRI as the cloning enzyme. The

library contains 55,296 clones with an average insert size

of 121 Kb providing approximately 15 haploid genome

equivalents. The deep coverage allows the isolation a

particular sequence with a probability of 99.9 %. Three

high density filters, each containing 18,432 clones

(doubly spotted), represent the whole library for colony

screening and can be requested from the Clemson University

BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match

3.8%; Score 41.4; DB 28; Length 765;



```
Best Local Similarity 53.4%; Pred. No. 9.6;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 560 CCGTCCAACTTGGGCTATCTCAGGACCGGAGAAAGTTGGGCGCACTGGCAAGCGTACG 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 CCATCCGGGTTACAGACTCGTGTGGAAGATGCTGACGTGGGCGCAGGCGCTCGCGCATGT 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 620 TTGCCACGAGGCGCAGACGACCGAGCGCGGAGCGACACCCCGACGCCCGTCACTGCAA 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 CGGCCACCGAGCCACCGCGCGGAGGAGGACACGACGAGCGGCGGCGGAGCTTGAGCGGCT 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 680 CCAGCGCGCTCCGAACTTGAAGCGGAAACACTTTACCTTTCCCTG 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 GCAGCGCCACCCACTTCCGCGTCCCTCTCTGCACCTTGCCTG 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
CNS06QTI 1053 bp DNA linear GSS 05-JUL-2001
LOCUS T7 end of clone AW0AA009A05 of library AW0AA from strain CLIB 89 of
Yarrowia lipolytica, genomic survey sequence.
ACCESSION AL411100
VERSION AL411100.1 GI:12180174
KEYWORDS GSS.
SOURCE Yarrowia lipolytica
ORGANISM Yarrowia lipolytica
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE
AUTHORS Souciet J.L., Aigle M., Artiguenave F., Blandin G.,
Bolotin-Fukuhara M., Bon E., Brottier P., Casaregola S.,
de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B.,
Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S.,
Saurin W., Tekai F., Toffano-Nioche C., Wesolowski-Louvel M.,
Wincker P. and Weissenbach J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 11152876
PUBMED 20584711
REFERENCE
AUTHORS Casaregola S., Neuveglise C., Lepingle A., Bon E., Feynertol C.,
Artiguenave F., Wincker P. and Gallardrin C.
TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
lipolytica
JOURNAL FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE 20584727
PUBMED 11152892
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
segrif@genoscope.cns.fr - Web : www.genoscope.cns.fr
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1. 1053
/organism="Yarrowia lipolytica"
/mol_type="genomic DNA"
/strain="CLIB 89"
/db_xref="taxon:4952"
/clone="AW0AA009A05"
/clone_lib="AW0AA"
/note="end : T7"

Best Local Similarity 53.4%; Pred. No. 9.6;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 560 CCGTCCAACTTGGGCTATCTCAGGACCGGAGAAAGTTGGGCGCACTGGCAAGCGTACG 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 CCATCCGGGTTACAGACTCGTGTGGAAGATGCTGACGTGGGCGCAGGCGCTCGCGCATGT 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 620 TTGCCACGAGGCGCAGACGACCGAGCGCGGAGCGACACCCCGACGCCCGTCACTGCAA 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 CGGCCACCGAGCCACCGCGCGGAGGAGGACACGACGAGCGGCGGCGGAGCTTGAGCGGCT 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 680 CCAGCGCGCTCCGAACTTGAAGCGGAAACACTTTACCTTTCCCTG 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 GCAGCGCCACCCACTTCCGCGTCCCTCTCTGCACCTTGCCTG 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
CNS06QTI 1053 bp DNA linear GSS 05-JUL-2001
LOCUS T7 end of clone AW0AA009A05 of library AW0AA from strain CLIB 89 of
Yarrowia lipolytica, genomic survey sequence.
ACCESSION AL411100
VERSION AL411100.1 GI:12180174
KEYWORDS GSS.
SOURCE Yarrowia lipolytica
ORGANISM Yarrowia lipolytica
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE
AUTHORS Souciet J.L., Aigle M., Artiguenave F., Blandin G.,
Bolotin-Fukuhara M., Bon E., Brottier P., Casaregola S.,
de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B.,
Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S.,
Saurin W., Tekai F., Toffano-Nioche C., Wesolowski-Louvel M.,
Wincker P. and Weissenbach J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 11152876
PUBMED 20584711
REFERENCE
AUTHORS Casaregola S., Neuveglise C., Lepingle A., Bon E., Feynertol C.,
Artiguenave F., Wincker P. and Gallardrin C.
TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
lipolytica
JOURNAL FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE 20584727
PUBMED 11152892
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
segrif@genoscope.cns.fr - Web : www.genoscope.cns.fr
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1. 1053
/organism="Yarrowia lipolytica"
/mol_type="genomic DNA"
/strain="CLIB 89"
/db_xref="taxon:4952"
/clone="AW0AA009A05"
/clone_lib="AW0AA"
/note="end : T7"

ORIGIN
Query Match 3.8%; Score 41.4; DB 29; Length 1053;
Best Local Similarity 56.1%; Pred. No. 11;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 611 AAGCGTACGTTGCCACGAGGCGCACGACGACCGCGGAGCGACACCCCGACGCCCG 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 858 AAACGGAGCGGCGACACGAGCTCCACGACTCCACCCCGGACGACACCAACCGCCAGCG 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 671 TCACTGCAACGAGGCGCTCGAACTTGAAGCGGAAACACTTTACCTTTCCCTGGCTAGAAA 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 ACATGACGAGCGAGCTACGAGCTGTACGCCACTGTACGCCACTGCCACCGCCAGGTGATG 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 731 ATGCGCTGGATCAATTACGA 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 GTGATGTCGATCAGAGCGA 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
CA704113 453 bp mRNA linear EST 26-NOV-2002
LOCUS wdkic.pk011.n7 wdkic Triticum aestivum cDNA clone wdkic.pk011.n7 5'
DEFINITION end, mRNA sequence.
ACCESSION CA704113
VERSION CA704113.1 GI:25425906
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
AUTHORS Tingey S.V., Powell W., Wolters P., Dolan M., Hainey C., Yuan Z.,
Miao G., Caraher N. and Hanafey M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
Location/Qualifiers
1. 453
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wdkic.pk011.n7"
/tissue_type="kernel"
/clone_lib="wdkic"
/note="Vector: pBluescript SK+; Wheat (Triticum aestivum
L.) developing kernel, 3 days after anthesis."

ORIGIN
Query Match 3.8%; Score 41.2; DB 14; Length 453;
Best Local Similarity 47.3%; Pred. No. 8.8;
Matches 156; Conservative 0; Mismatches 173; Indels 1; Gaps 1;

Qy 440 ACGACACCAAGAACCCATTGACGCTTCGCGATCCAACTATCGGTATCAATTCGGA 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 ACCAAGCATATAACCTACACACGACCAAGTAATACCAAGCTA-CAAATCGGCT 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 500 ACACGCGCGGACTCGGGACTCTATTCCAGGCTTGTGTGCGACCTTCGGATTACCTA 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 AAACAGAGCACACACGGAACATGAATGATGATGAAGATCGAATCGAACCGGCCACCTCCA 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 560 CCGTCCAACTTGGAGGCTTATCTCAGGACCGGAGGAAAGTTGGCGCAACTGGCAAGCGTACG 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 TATAGATTGATCAGCTGTCTTGAACATGGTGACGCTGCGCGCGCTGGAGCATGT 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



[illegible]

**AUTHORS** Kirkness,E.P., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,  
Rusch,D.B., Deichez,A.L., Pop,M., Wang,W., Fraser,C.M. and  
Venter,J.C.  
**TITLE** The dog genome: survey sequencing and comparative analysis  
**JOURNAL** Science 301 (5641), 1898-1903 (2003)  
**MEDLINE** 22875432  
**PUBMED** 14512627  
**COMMENT** Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirkness@tigr.org  
Class: shotgun.  
**FEATURES** Location/Qualifiers  
source  
1..551  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/notes="Site 1: BstXI; Libraries were prepared from  
peripheral blood"  
**ORIGIN**  
Query Match 3.6%; Score 39.4; DB 29; Length 551;  
Best Local Similarity 52.8%; Pred. No. 28;  
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 454 CCATTACAGCTCTCGGATCCAACTATCGTGTATCAATTCGGAACCGCGCGACT 513  
Db 14 CACCTGACCTCTCTCGAACCCTGGTGACTCTCTTTCCACATGGAACTCATCTGATC 73  
QY 514 CGCGACTCTATTCCAAAGGCTTGTGTGCACCTTCGGATTACCTACCGTCCAACTTGAG 573  
Db 74 TGGTGACTCTGGTCTCAGTTGGCTCACTCGACCTGGTGCTCTCTCGTCCACACTTGG 133  
QY 574 GCCTATCTACGACCGAGGAAGTGTGGCAACTGGCAAGC 614  
Db 134 ACCTAACTCAGAACTGTGAATCTTGCCACCTAGGCCGC 174  
**RESULT 13.**  
CA255506 577 bp mRNA linear EST 26-SEP-2003  
LOCUS SCEPFL4180E09.g FL4 Saccharum officinarum cDNA clone SCEPFL4180E09  
5', mRNA sequence.  
ACCESSION CA255506  
VERSION CA255506.1 GI:35940463  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Saccharum.  
1 (bases 1 to 577)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenhariaia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bccccenter.fcav.unesp.br  
Plate: 180 row: E column: 09  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
**FEATURES**  
source  
1..577  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCEPFL4180E09"  
/lab\_host="DH10B"  
/clone\_lib="FL4"  
/note="Organ: Developed inflorescence and rachis  
(20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;  
An unidirectional cDNA library generated from [Developed  
inflorescence and rachis (20cm-long)]. cDNA was prepared  
from polyA+ mRNA using SuperScript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucst.lad.ic.unicamp.br/public"  
**ORIGIN**  
Query Match 3.6%; Score 38.6; DB 14; Length 577;  
Best Local Similarity 58.1%; Pred. No. 47;  
Matches 68; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 587 CCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGCCAGGAGCCACGACGACGCG 646  
Db 537 CTTTGGAGAGTTGCGGGAACCGCGCGCGGTGTCACCCCGTAGCCACGGGCACAG 478  
QY 647 CCGAGGCGCAACCCCGGACGCGCTACTGCAACGAGCGCTCCGAATTTGAGCGG 703  
Db 477 CGCGCGTAGCTCGCGCCCTCGCGCGCTGTAGCCACGCGCGCGCTAGCGCGCGCG 421  
**RESULT 14**  
CF877969 717 bp mRNA linear EST 31-OCT-2003  
LOCUS tric077xp18.bl T.reesei mycelial culture, Version 6 October 2003  
DEFINITION Hypocrea jecorina cDNA clone tric077xp18, mRNA sequence.  
ACCESSION CF877969  
VERSION CF877969.1 GI:38132651  
KEYWORDS EST.  
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
ORGANISM Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.  
REFERENCE 1 (bases 1 to 717)  
AUTHORS Diener,S.E., Dankmeyer,L., Dunn-Coleman,N., Houfek,T.D.,  
Mitchell,T.K., van Solingen,P., Teunissen,P.J.M., Ward,M. and  
Dean,R.A.  
TITLE Analysis of the protein processing and secretion pathways in a  
Trichoderma reesei EST dataset  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA  
Tel: 919-513-0020  
Fax: 919-513-0024  
Email: ralph.dean@ncsu.edu  
Seq primer: Tr-F1 primer.  
Location/Qualifiers  
**FEATURES** source  
1..717  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="QM6a"  
/db\_xref="taxon:51453"  
/clone="tric077xp18"  
/dev\_stage="mycelia"  
/clone\_lib="T.reesei mycelial culture, Version 6 October  
2003"  
/note="Vector: pREP3V, Site 1: Not I/Sal I; Mycelial  
culture grown from 24 hrs to 6 days with varying Carbon  
and Nitrogen sources and concentrations."

## ORIGIN

Query Match 3.6%; Score 38.6; DB 14; Length 717;  
Best Local Similarity 49.3%; Pred. No. 51;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 585 GACCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGGCAGCGAGGCCACGACGACGAG 644  
Db 488 GAAGAAGGCCAAGAAGGCGCAAGCGAAGGCGTTCAAGGGCTCCGTCGCCAACATGTCCCT 547  
Qy 645 CGCGGAGCGACACCCCGCGCGCTCACTGCAACCGCGCTCGGAACCTTGAAGCGGA 704  
Db 548 CGGTGGTGGCAAGACCCAGCGCTTTGACGCTCGCGTCAACGCGCGCTCGCGCGCGGT 607  
Qy 705 ACATTTTACCTTCCCTGGCTAGAAATGGCGTGGATCAATTAGAACCGACACCCCGCAA 764  
Db 608 CCATTTTGGCTTGGCTCGCGCAACGACACGCTGATGTTGCAACTACTCCCGCGTGC 667

Qy 765 CGAAATTCAAAGCTTACTTCCGT 789

Db 668 CGCCACTGAGCCCTCACCGTCGT 692

## RESULT 15

CF865608

LOCUS 718 bp mRNA linear EST 31-OCT-2003  
DEFINITION trico02xel17.b1 T.reesei mycelial culture, version 6 October 2003  
Hypocrea jecorina cDNA clone trico02xel17, mRNA sequence.

CF865608

CF865608.1 GI:38120234

EST.

SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)

ORGANISM

Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

## REFERENCE

1 (bases 1 to 718)  
Diener S.E., Dankmeyer L., Dunn-Coleman N., Houfek T.D.,  
Mitchell T.K., van Solingen P., Teunissen P.J.M., Ward M. and  
Dean R.A.

TITLE Analysis of the protein processing and secretion pathways in a  
Trichoderma reesei EST dataset

JOURNAL

Unpublished (2003)

COMMENT

Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA

Tel: 919-513-0020

Fax: 919-513-0024

Email: ralph\_dean@ncsu.edu

Seq primer: LT-F1 primer.

Location/Qualifiers

FEATURES

source

1..718  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="QM6a"  
/db\_xref="taxon:51453"  
/clone="trico02xel17"  
/dev\_stage="mycelia"  
/clone\_lib="T.reesei mycelial culture, Version 6 October 2003"  
/note="Vector: pREP3Y; Site1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."

## ORIGIN

Query Match 3.6%; Score 38.6; DB 14; Length 718;  
Best Local Similarity 49.3%; Pred. No. 51;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 585 GACCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGAGGCCACGACGACGAG 644

Db 254 GAAGAAGGCCAAGAAGGCGCAAGCGGAGGGCTTCAAGGGCTCCGTCGCCAACATGTCCCT 313

Qy 645 CGCGGAGCGACACCCCGCGCGCTCACTGCAACCGAGCGCTCCGAACTTGAAGCGGA 704  
Db 314 CGGTGGTGGCAAGACCCAGGCTCTTGACGCTGCCGTCAACGCCGCCCTCCGCGCGGTGT 373  
Qy 705 ACATTTTACCTTCCCTGGCTAGAAATGGCGTGGATCAATTACGAACCGACACCCCGCAA 764  
Db 374 CCATTTTGGCGTGGCTGCCGCAACGACGCTGATGTTGCAACTACTTCCCGCGTGC 433  
Qy 765 CGAAATTCAAAGCTTACTTCCGT 789  
Db 434 CGCCACTGAGCCCTCACCGTCGT 458

Search completed: March 8, 2004, 01:19:36  
Job time : 3281 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 01:19:45 ; Search time 129 Seconds  
(without alignments)  
792.885 Million cell updates/sec

Title: US-09-993-777-7  
Perfect score: 1863  
Sequence: 1 MASLGLTALLAATLAPFGA.....VELVAIVNPSALSPPDSIKM 362

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	362	2 AAW00635	AAW00635 ILTV glyc
2	1863	100.0	362	2 AAW06787	AAW06787 ILTV glyc
3	1863	100.0	362	5 ABG31711	ABG31711 Infectiou
4	156	8.4	384	2 AAR47236	AAR47236 Wild-type
5	152	8.2	355	2 AAW11475	AAW11475 Marek's d
6	152	8.2	355	2 AAW13105	AAW13105 Marek's d
7	148	7.9	364	2 AAW22999	AAW22999 Canine he
8	148	7.9	364	2 AAW72659	AAW72659 Canine he
9	148	7.9	364	4 AAB51316	AAB51316 Canine he
10	148	7.9	364	7 AAE39133	AAE39133 CHV PCG13
11	143.5	7.7	350	1 AAP70645	AAP70645 Pseudorab
12	143.5	7.7	350	2 AAR63143	AAR63143 Glycoprot
13	143.5	7.7	350	4 AAE05396	AAE05396 Pseudorab
14	143	7.7	317	2 AAR80637	AAR80637 Bovine he
15	142.5	7.6	350	4 AAU04970	AAU04970 Pseudorab
16	141.5	7.6	350	4 AAB82502	AAB82502 Pseudorab
17	140.5	7.5	380	2 AAR48063	AAR48063 Sequence
18	139	7.5	371	4 AAB78825	ABB78825 Pseudorab
19	138	7.4	370	3 AAY32474	AAY32474 DNA encod
20	125	6.7	356	2 AAW03136	AAW03136 Herpesvir
21	125	6.7	356	2 AAW03323	AAW03323 Herpesvir
22	123.5	6.6	253	2 AAR22323	AAR22323 Marek Dis
23	119	6.4	423	7 ADC42841	ADC42841 REMAP pro
24	115	6.2	612	6 AAE33343	AAE33343 Mouse CD4
25	114	6.1	456	2 AAR33180	AAR33180 L3T4 muta

## ALIGNMENTS

## RESULT 1

AAW00635  
ID AAW00635 standard; protein; 362 AA.

XX AAW00635;

XX 16-OCT-2003 (revised)

DT 19-NOV-1996 (first entry)

XX ILTV glycoprotein gi.

XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;  
vector; vaccine; chicken; poultry; immunisation; glycoprotein gi.

XX Gallid herpesvirus 1.

PN W09508622-A1.

XX 30-MAR-1995.

XX 16-SEP-1994; 94WO-US010628.

XX 24-SEP-1993; 93US-00126597.

XX (SYTR ) SYNTRO CORP.

XX Wild MA, Cochran MD;

XX WPI; 1995-139591/18.

XX N-PSDB; AAT33504.

XX Recombinant attenuated infectious laryngotracheitis virus - for use in  
vaccines to protect poultry from infection from the virus, also methods  
of distinguishing between vaccinated and naturally infected birds.

XX Example 1; Page 102-103; 177pp; English.

XX The gi gene, spanning nucleotides 9874-10962 of the unique short region  
(AAT33504) of infectious laryngotracheitis virus (ILTV), codes for a  
glycoprotein (AAW00635) of approx. 39,7535 mol.wt. The gi glycoprotein is  
homologous to Varicella-Zoster gi. Deletion of the gi gene results in an  
attenuated ILTV that is useful as a vaccine against ILTV disease in  
chickens. Recombinant virus deleted for gi was safe in animal trials.  
Deletion of the gi gene serves as a negative marker to distinguish  
vaccines from infected animals. A gene coding for a foreign antigen may  
be inserted into the gi gene to produce a recombinant multivalent  
vaccine. (Updated on 16-OCT-2003 to standardise OS field)

XX

Aay08621 Human sec  
Aay67312 Human sec  
Adc38775 Human sec  
Aar33181 L3T4 muta  
Abb69348 Drosophil  
Aar95878 MD-40 dom  
Aar53522 Protein s  
Abp70863 Yeast TUP  
Aaw72164 HSV-2 str  
Aaw72012 HSV-2 str  
Aaw72144 HSV-2 str  
Aab70447 Amino aci  
Aam9027 Human pol  
Ada25189 CD4 epit  
Abb58158 Drosophil  
Aab70446 Amino aci  
Aay25768 Human sec  
Abb58595 Drosophil  
Abp33907 Human ORF  
Abu32202 Protein e

26 112 6.0 387 2 AAY08621  
27 112 6.0 387 3 AAY67312  
28 112 6.0 387 7 ADC38775  
29 109.5 5.9 454 2 AAR33181  
30 108.5 5.8 319 4 ABB69348  
31 108 5.8 713 2 AAR95878  
32 108 5.8 713 6 ABR53522  
33 108 5.8 713 6 ABP70863  
34 107 5.7 372 2 AAW72164  
35 107 5.7 433 2 AAW72012  
36 107 5.7 433 2 AAW72144  
37 106.5 5.7 378 6 ABP70447  
38 105.5 5.7 378 4 AAM9027  
39 105 5.6 432 6 ADA25189  
40 103 5.5 982 4 ABB58158  
41 102.5 5.5 378 6 ABP70446  
42 102.5 5.5 379 2 AAY25768  
43 102.5 5.5 4498 4 ABB58595  
44 102 5.5 234 5 ABP33907  
45 99 5.3 961 6 ABU32202

SQ Sequence 362 AA;

Query Match 100.0%; Score 1863; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-173;  
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATTIQLQLFNPGQ 60  
 DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATTIQLQLFNPGQ 60

QY 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCNTHRSSVVFVGCCKTETVTFASNRLTG 120  
 DB 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCNTHRSSVVFVGCCKTETVTFASNRLTG 120

QY 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180  
 DB 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RPFGLPTVQLEAYLRTEESWRNMQAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240  
 DB 181 RPFGLPTVQLEAYLRTEESWRNMQAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVVAAVSATIGLVIVISIVTRNMCTP 300  
 DB 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVVAAVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIYNPSALSPDSI 360  
 DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIYNPSALSPDSI 360

QY 361 KM 362  
 DB 361 KM 362

RESULT 2  
 AAW06787  
 ID AAW06787 standard; protein; 362 AA.  
 XX AC AAW06787;  
 XX DT 16-OCT-2003 (revised)  
 XX DT 02-JUN-1997 (first entry)  
 XX DE ILTV glycoprotein gI.  
 KW ILTV; vaccine; vector; attenuation; poultry;  
 KW avian infectious bronchitis virus; Newcastle disease virus;  
 KW infectious bursal disease virus of chickens; Marek's disease virus;  
 KW herpesvirus; glycoprotein gI.  
 XX OS Gallid herpesvirus 1; USDA strain 8302.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..22  
 XX FT /label= Sig\_peptide  
 XX FT Protein 23..362  
 XX FT /label= Mat\_protein  
 XX FT Region 272..292  
 XX FT /label= Transmembrane\_helix  
 XX PN WO9629396-A1.  
 XX XX 26-SEP-1996.  
 XX XX 21-MAR-1996; 96WO-US003916.  
 XX PR 23-MAR-1995; 95US-00410121.  
 XX PR 06-JUN-1995; 95US-00468190.  
 XX PA (SYTR) SYNTRO CORP.  
 XX XX Wild MA, Cochran MD;  
 PI

XX WPI; 1996-443172/44.  
 DR N-PSDB; AAT44384, AAT44385.  
 XX Recombinant infectious laryngotracheitis virus with deletion in the  
 PT glycoprotein G, gI or US2 gene, etc. - useful for vaccines against  
 PT infectious laryngotracheitis in poultry.  
 XX Example 11; Page 110-111; 216pp; English.  
 XX Glycoprotein gI (AAW06787) is encoded by ORF8 of the unique short region  
 CC (AAT44384) of infectious laryngotracheitis virus (ILTV). It shows  
 CC homology to the varicella zoster virus gI glycoprotein. Recombinant ILTV  
 CC gI protein produced in a swinepox virus reacts to convalescent sera from  
 CC ILTV-infected chickens. Deletion of the gI gene results in an attenuated  
 CC ILTV that is useful as a vaccine and as a negative marker to distinguish  
 CC vaccinated from infected animals. Insertion of a foreign gene into the gI  
 CC gene allows prodn. of multivalent vaccines. (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 XX SQ Sequence 362 AA;

Query Match 100.0%; Score 1863; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-173;  
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATTIQLQLFNPGQ 60  
 DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATTIQLQLFNPGQ 60

QY 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCNTHRSSVVFVGCCKTETVTFASNRLTG 120  
 DB 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCNTHRSSVVFVGCCKTETVTFASNRLTG 120

QY 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180  
 DB 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RPFGLPTVQLEAYLRTEESWRNMQAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240  
 DB 181 RPFGLPTVQLEAYLRTEESWRNMQAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVVAAVSATIGLVIVISIVTRNMCTP 300  
 DB 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVVAAVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIYNPSALSPDSI 360  
 DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIYNPSALSPDSI 360

QY 361 KM 362  
 DB 361 KM 362

RESULT 3  
 ABG31711  
 ID ABG31711 standard; protein; 362 AA.  
 XX AC ABG31711;  
 XX XX 29-AUG-2003 (revised)  
 XX DT 15-NOV-2002 (first entry)  
 XX XX Infectious laryngotracheitis virus (ILTV) glycoprotein I.  
 XX KW Infectious laryngotracheitis virus; novel recombinant avian herpesvirus;  
 KW NAHV; herpes virus of turkeys; HVT; Marek's disease; Newcastle disease;  
 KW infectious laryngotracheitis; virucide; immunostimulant; vaccine; ILTV;  
 XX glycoprotein I.  
 XX OS Gallid herpesvirus 1.  
 XX PI

PN	US2002081316-A1.	
XX		
PD	27-JUN-2002.	
XX		
PF	14-JUN-2001; 2001US-00881457.	
XX		
PR	12-JUN-1992; 92US-00898087.	
PR	26-FEB-1993; 93US-00023610.	
PR	14-JUN-1993; 93WO-US0005681.	
PR	09-AUG-1994; 94US-00288065.	
PR	09-AUG-1995; 95WO-US010245.	
PR	13-JUN-1996; 96US-00663566.	
PR	21-FEB-1997; 97US-00804372.	
PR	23-OCT-1999; 99US-00426352.	
XX		
PA	(COCH/) COCHRAN M D.	
PA	(COOK/) COOK S M.	
PA	(WILD/) WILD M A.	
XX		
PI	Cochran MD, Cook SM, Wild MA;	
XX		
DR	WPI; 2002-635456/68.	
DR	N-PSDB; ABK90557.	
XX		
PT	Novel recombinant avian herpesvirus comprising unique long and repeat	
PT	viral genome regions of herpes virus of turkeys, unique short viral	
PT	genome region of Marek's disease virus, and optional foreign DNA	
PT	sequence.	
XX		
PS	Disclosure; Page 21-22; 26pp; English.	
XX		
CC	The invention relates to a novel recombinant avian herpesvirus (NAHV)	
CC	comprising a herpes virus of turkeys (HVT) unique long (UL) and repeat	
CC	viral genome region and a Marek's disease virus unique short (US) viral	
CC	genome region where at least one foreign DNA sequence is inserted within	
CC	a US2 gene of the US region of the NAHV, and is capable of being	
CC	expressed in a host cell. NAHV is useful for producing vaccines used for	
CC	immunising an avian species against Newcastle disease, infectious	
CC	laryngotracheitis and Marek's disease. This sequence represents	
CC	infectious laryngotracheitis virus (ILT) glycoprotein I. (Updated on 29-	
CC	AUG-2003 to standardise OS field)	
XX		
SQ	Sequence 362 AA;	
	Query Match 100.0%; Score 1863; DB 5; Length 362;	
	Best Local Similarity 100.0%; Pred. No. 1.5e-173;	
	Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MASLLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPENTATLQLFPMGQ 60	
Db	1 MASLLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPENTATLQLFPMGQ 60	
Qy	61 RPHKPSGTVVAFRSITNOCYQELSEERFENCNTHSSSVFVGCCKVTEYFASNRLTG 120	
Db	61 RPHKPSGTVVAFRSITNOCYQELSEERFENCNTHSSSVFVGCCKVTEYFASNRLTG 120	
Qy	121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180	
Db	121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180	
Qy	181 RTFGLPTVQLBAYLRTEESWRNWAQYVATEATTTSAEATTPPTVATASASELAHEFTFP 240	
Db	181 RTFGLPTVQLBAYLRTEESWRNWAQYVATEATTTSAEATTPPTVATASASELAHEFTFP 240	
Qy	241 WLENGVDHYETTPANENSNVRLTGMSPTLIGTVAAVVSATIGLVISIVTRNNCTP 300	
Db	241 WLENGVDHYETTPANENSNVRLTGMSPTLIGTVAAVVSATIGLVISIVTRNNCTP 300	
Qy	301 HRKLDTVSDDEERSQTRRSKFGPMVACINKGADQDSSELVELVAIVNPSALSSPDSI 360	
Db	301 HRKLDTVSDDEERSQTRRSKFGPMVACINKGADQDSSELVELVAIVNPSALSSPDSI 360	
Qy	361 KM 362	

Db	361 KM 362	
RESULT 4		
AAAR47236		
ID	AAAR47236 standard; protein; 384 AA.	
XX		
AC	AAAR47236;	
XX		
DT	27-AUG-2003 (revised)	
DT	25-MAR-2003 (revised)	
DT	06-SEP-1994 (first entry)	
XX		
DE	Wild-type Feline Herpes Virus ORF-1-encoded protein.	
XX		
KW	Feline herpes virus; FHV; genome; mutant; vaccine; ORF-1;	
KW	feline viral rhinotracheitis; herpesviridae family;	
KW	herpes virus A subfamily.	
XX		
OS	Feline herpesvirus.	
XX		
PN	WO9403621-A1.	
XX		
PD	17-FEB-1994.	
PF	23-JUL-1993; 93WO-EP001971.	
XX		
PR	30-JUL-1992; 92EP-00202365.	
XX		
PA	(ALKU ) AKZO NV.	
XX		
PI	Sondermeijer PJA, Willemse MJ;	
XX		
DR	WPI: 1994-065709/08.	
DR	N-PSDB; AAQ56188.	
XX		
PT	Feline herpes virus mutant comprising a heterologous gene inserted in the	
PT	virus genome - for vaccination against feline herpes virus and other	
PT	feline pathogens.	
XX		
PS	Claim 2; Page 37-38; 55pp; English.	
XX		
CC	Mutant versions of the wild-type feline herpes virus genome (AAQ56188)	
CC	are claimed. Esp. the FHV mutant is one which fails to produce one or	
CC	more antigenic or functional polypeptides. The mutant may contain a	
CC	heterologous nucleic acid sequence inserted within one of the 6 open	
CC	reading frames. The attenuated FHV mutants can be used to prepare	
CC	vaccines against feline viral rhinotracheitis and, where the heterologous	
CC	insert encodes an antigen of a feline pathogen, against other infectious	
CC	diseases of felidae. (Updated on 25-MAR-2003 to correct PN field.)	
CC	(Updated on 27-AUG-2003 to correct OS field.)	
XX		
SQ	Sequence 384 AA;	
	Query Match 8.4%; Score 156; DB 2; Length 384;	
	Best Local Similarity 23.3%; Pred. No. 2.9e-06;	
	Matches 75; Conservative 48; Mismatches 141; Indels 58; Gaps 12;	
Qy	5 LGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRP-ATIQQLFFPMGQ-RP 62	
Db	1 MSSIAPIYILMAITGVYGVIRGDHVLVDTSFGVIYPTLENFTIYGHLLIFLDDQPLP 60	
Qy	63 HRPYSGTVVAFRSITNOCYQELSEERFENCNTHSSSVFVGC--KVTEYFASNRLTG 120	
Db	61 VNNYNGTLEI-IHYNHSSCYKIVQVIEYSSCFVRNNAFRSCLHKTSHQDDQSINTS 119	
Qy	121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180	
Db	120 VETGMLLTITSPKMDGGIYALVRFRFNHNNK-ADVFGLSVFVYSP----DTGHRHAD- 173	
Qy	181 RTFGLPTVQLBAYLRTEESWRNWAQYVATEATTTSAEATTPPTVATASASELAHEFTFP 240	



Db 174 -----ENLGEILTPS--SMETVVKVNTPIYDHMTTQT-----TSKMSHSE-----P 216

Qy 241 WLENGVDHYEPTPANENSVTVRL-----GTMSPTLIGTVVAAVVSAT- 283

Db 217 SNTSISCHTFQNDPNEGELYTHLNIAGNIYYDDWMDGTTLQRLIDMGLNLSVTSSP 276

Qy 284 -----IGLVIVISI 292

Db 277 KNETTQKWTDPDRKVGIVIS 298

RESULT 5

AAW11475

ID AAW11475 standard; protein; 355 AA.

XX

AC AAW11475;

XX

DT 30-APR-1997 (first entry)

XX

DE Marek's disease virus glycoprotein gI.

XX

XX MDV; UL32; membrane glycoprotein 82; glycoprotein gI; antigen; vaccine;

KW vector; fowlpox virus.

XX

OS Marek's disease gammaherpesvirus type 1 strain GA.

XX

PN WO9703187-A2.

XX

PD 30-JAN-1997.

XX

PF 05-JUL-1996; 96WO-US011360.

XX

XX 07-JUL-1995; 95US-00499474.

XX

XX (JAPG ) NIPPON ZEON KK.

PA (USDA ) US SEC OF AGRIC.

PA

XX Lee LF, Nazerian K, Witter RL, Wu P, Yanagida N, Yoshida S;

PI WPI; 1997-119044/11.

XX

XX New DNA encoding glyco:protein 82 of Marek disease virus - useful in

PT vaccines to protect poultry.

PT

PS Disclosure; Page 72-73; 101pp; English.

XX

XX Genes encoding Marek's disease virus glycoprotein gI (AAW11475) and/or

CC glycoprotein gE (AAW11474) can be incorporated into recombinant viral

CC vectors that also carry the UL32 gene (see also AAT51358) encoding

CC glycoprotein 82 (AAW11473). The gI and gE genes were cloned from MDV

CC genomic DNA by PCR (see also AAT51368-69 and AAT51366-67). A transfer

CC vector was constructed that was used to produce a recombinant fowlpox

CC virus useful as a vaccine to protect poultry against MDV infection

XX

SQ Sequence 355 AA;

Query Match 8.2%; Score 152; DB 2; Length 355;

Best Local Similarity 22.9%; Pred. No. 6.5e-06;

Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;

Qy 19 GAMGIVITGNHVSARIDDDHIVIVAPPE-ATIQQLFFMPGQRPKPKYSGTVRVAFRSD 77

Db 15 GIWSIVTGTSTVLSTQDSALVAFGLDKMNVVRGQLFLGDTQRTSSYTGTTET-LKWD 73

Qy 78 ITNOCYQELSERFENCNTHRSSSVFVGCKVTEYTFASNRLTGPPHPEK---LTIRNRP 134

Db 74 BEYKCYSLHATSYMDCPAIDATVFRGCRDAVVAQAQPHDRV--QPFPEKGTLLRIVEPRV 131

Qy 135 NDSCGMFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASCRFTGLPTVQLEAYL 194

Db 132 SDTGSYYIRVALAG-RNMSDIFRMAVIRS-----SKSWACHNSASSFOAHKCI 179

Qy 195 RTEE--SWRNQAYVATEATTTSABATTPVTATTSASELEAEHFTFPWLENGVDHYEPT 252

Db 180 RYDRAFAFENYLIGHVGNLLDSSELHAYNITPQIS-TDINIITPPFYDNGSTIYSPT 238

Qy 253 P---ANENSNTVRLGT-XSPTLIGTVVAAV--SATIGLVIV---ISIVTRNMCTPHRK 303

Db 239 VENLFNNSHVDAMNSGMMNTVLKVTLPRLIYFSTMIVLCTIALAIYLVCCRCRSPHER 298

Qy 304 LDTVSQDDERSQTRRESRKFGPMVACEINKGADQDSELVE 344

Db 299 I---YICEPRSD-----APLITSAVNESFQYDYNVKE 328

RESULT 6

AAW13105

ID AAW13105 standard; protein; 355 AA.

XX

AC AAW13105;

XX

DT 17-OCT-2003 (revised)

DT 12-MAY-1997 (first entry)

XX

DE Marek's disease virus type 1 glycoprotein I.

XX

XX GA strain; glycoprotein; gPI; gpE; recombinant virus;

KW Newcastle disease virus; herpes virus; vaccine.

XX

OS Marek disease virus type 1.

XX

PN JP09009978-A.

XX

PD 14-JAN-1997.

XX

PF 26-APR-1996; 96JP-00131084.

XX

XX 28-APR-1995; 95JP-00129523.

XX

XX (JAPG ) JAPANESE GEON CO LTD.

XX

XX WPI; 1997-126432/12.

DR N-PSDB; AAT61883.

DR

XX Recombinant virus contg. herpes virus glyco:protein gE - and effective

PT Newcastle disease vaccine contg. it.

PT

XX Claim 5; Page 14-15; 20pp; Japanese.

XX

XX The present sequence encodes the Marek's disease virus (MDV) type 1 (GA

CC strain) glycoprotein (gp) I. A recombinant virus containing the gene

CC encoding MDV gpE, and optionally also gpI, in a genomic region

CC unessential for its growth, can be used as the effective component in a

CC Newcastle disease virus vaccine. (Updated on 17-OCT-2003 to standardise

CC OS field)

XX

SQ Sequence 355 AA;

Query Match 8.2%; Score 152; DB 2; Length 355;

Best Local Similarity 22.9%; Pred. No. 6.5e-06;

Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;

Qy 19 GAMGIVITGNHVSARIDDDHIVIVAPPE-ATIQQLFFMPGQRPKPKYSGTVRVAFRSD 77

Db 15 GIWSIVTGTSTVLSTQDSALVAFGLDKMNVVRGQLFLGDTQRTSSYTGTTET-LKWD 73

Qy 78 ITNOCYQELSERFENCNTHRSSSVFVGCKVTEYTFASNRLTGPPHPEK---LTIRNRP 134

Db 74 BEYKCYSLHATSYMDCPAIDATVFRGCRDAVVAQAQPHDRV--QPFPEKGTLLRIVEPRV 131

Qy 135 NDSCGMFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASCRFTGLPTVQLEAYL 194

Db 132 SDTGSYYIRVALAG-RNMSDIFRMAVIRS-----SKSWACHNSASSFOAHKCI 179

Qy 195 RTEE--SWRNQAYVATEATTTSABATTPVTATTSASELEAEHFTFPWLENGVDHYEPT 252

Db 180 RYVDRMAFENYLIGHVGNLSDSELHAIYNITPQIS-TDINIITPPYDNGSTIYSPT 238  
QY 253 P-ANENSNTVRLGT-MSPTLIGTVAAV--SATIGLVIV--ISIVTRNMCPTREK 303  
Db 239 VFNLFNNSHVDAMNSTMWNTVLKYLPLRIYFTSWIVLCIITAIYLVCEKCRSPHRR 298  
QY 304 LDTVSQDDEBSQTRRESKFGPMVACEINKGADQDSSELVE 344  
Db 299 I-----YIGERSDE-----APLITSAVNESFYDYNVKE 328

RESULT 7  
AAW22999  
ID AAW22999 standard; protein; 364 AA.  
XX  
AC AAW22999;  
XX  
DT 20-FEB-1998 (first entry)  
XX  
DE Canine herpesvirus glycoprotein I PCgI364.  
XX  
KW Vaccine; vector; gene therapy; canid; dog; CHV; CgI; PCgI364;  
KW glycoprotein I.  
XX  
OS Canine herpesvirus.  
XX  
WO929772-A1.  
PN  
XX  
PD 21-AUG-1997.  
XX  
PF 14-FEB-1997; 97WO-US004115.  
XX  
PR 15-FEB-1996; 96US-00602010.  
XX  
PA (HESK-) HESKA CORP.  
XX  
PI Haanes EU, Frank RA;  
XX  
DR WPI; 1997-424758/39.  
DR N-PSDB; AAT75616.  
XX  
PT Recombinant canine herpes virus and its genome - useful as vaccine to  
PT protect canids against infectious, metabolic or genetic diseases.  
XX  
PS Claim 57; Page 181-182; 240pp; English.  
XX  
CC This protein comprises canine herpesvirus (CHV) glycoprotein I (PCgI364).  
CC Its sequence was deduced from a coding region found in CHV genomic DNA  
CC molecule NCUS10592 (see AAT75616). PCgI364 can be expressed in  
CC transformed host cells. CHV proteins, nucleic acids, and antibodies  
CC raised against CHV proteins, can be used to protect canids against CHV  
CC infection. Novel recombinant CHV and novel recombinant CHV genomes that  
CC contain heterologous nucleic acid molecules inserted e.g. into CgI gene,  
CC can be used as vaccines to protect canids against infectious, metabolic  
CC or genetic diseases  
XX  
SQ Sequence 364 AA;

Query Match 7.9%; Score 148; DB 2; Length 364;  
Best Local Similarity 27.3%; Pred. No. 1.7e-05;  
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;  
QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQFFMPQQR-P 62  
Db 16 LLITMFLPILFLFLYGVNGFYKGTIYSMFLNTSSGFSIFPDDKFIIVSGRLFLDDQHL 75  
QY 63 HKPYSGTVRVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFASNRLTG 120  
Db 76 VNNYSGTIEFTI---HFNNSCTVYTQTIETIFSCPIFNNAFRSCLKVKVSKHESQLRINS 132  
QY 121 PPHPFKLTIRNPRNDSGMFVIVRLDDTKPEIDVFAIQLSVYQF 165  
Db 133 IENGVLLEITNPKNDSGVFIIRVQLENNK--TDVFGIPAFIYSF 175

RESULT 9  
AAB51316  
ID AAB51316 standard; protein; 364 AA.  
XX  
AC AAB51316;  
XX  
DT 06-AUG-2003 (revised)

RESULT 8  
AAW72659  
ID AAW72659 standard; protein; 364 AA.  
XX  
AC AAW72659;  
XX  
DT 27-AUG-2003 (revised)  
DT 07-JAN-1999 (first entry)  
XX  
DE Canine herpes virus protein sequence PCgI-364.  
XX  
KW Canine herpes virus; CHV; recombinant canine herpes virus vector; genome;  
KW vaccination; dog; protozoan; helminth; ectoparasite; bacteria;  
KW virus infection.  
XX  
OS Canine herpesvirus.  
XX  
US5804197-A.  
PN  
XX  
PD 08-SEP-1998.  
XX  
PF 12-JUL-1996; 96US-00680726.  
XX  
PR 15-FEB-1996; 96US-00602010.  
XX  
PA (HESK-) HESKA CORP.  
XX  
PI Frank RS, Haanes EU;  
XX  
DR WPI; 1998-505590/43.  
DR N-PSDB; AAV66941.  
XX  
CC Canine herpes virus nucleic acids - useful for producing recombinant  
CC canine herpes virus vectors.  
XX  
PS Claim 7; Col 133-136; 103pp; English.  
XX  
CC The present sequence represents a canine herpes virus (CHV) protein  
CC sequence. The nucleic acid sequence which encodes the protein can be used  
CC for DNA vaccination of dogs against CHV and also other infective agents  
CC such as protozoans, helminths, ectoparasites, bacteria and viruses. CHV  
CC can be formulated by incorporation of heterologous nucleic acid molecules  
CC as a single multivalent therapeutic composition against a variety of  
CC canine pathogens. (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 364 AA;

Query Match 7.9%; Score 148; DB 2; Length 364;  
Best Local Similarity 27.3%; Pred. No. 1.7e-05;  
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;  
QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQFFMPQQR-P 62  
Db 16 LLITMFLPILFLFLYGVNGFYKGTIYSMFLNTSSGFSIFPDDKFIIVSGRLFLDDQHL 75  
QY 63 HKPYSGTVRVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFASNRLTG 120  
Db 76 VNNYSGTIEFTI---HFNNSCTVYTQTIETIFSCPIFNNAFRSCLKVKVSKHESQLRINS 132  
QY 121 PPHPFKLTIRNPRNDSGMFVIVRLDDTKPEIDVFAIQLSVYQF 165  
Db 133 IENGVLLEITNPKNDSGVFIIRVQLENNK--TDVFGIPAFIYSF 175

RESULT 9  
AAB51316  
ID AAB51316 standard; protein; 364 AA.  
XX  
AC AAB51316;  
XX  
DT 06-AUG-2003 (revised)



```

XX AAP70645;
AC
XX 29-APR-1991 (first entry)
DT
XX
XX Pseudorabies virus gp63 protein.
DE
XX PRV; vaccine; gI; gp50; gp63.
KW
XX Pseudorabies virus.
OS
XX W08702058-A.
PN
XX
XX 09-APR-1987.
PD
XX
XX 28-AUG-1986; 86WO-US001761.
PF
XX
XX 04-OCT-1985; 85US-00784787.
PR
XX 26-NOV-1985; 85US-00801799.
PR
XX 26-MAR-1986; 86US-00844113.
PR
XX 26-MAR-1986; 86US-00844133.
PR
XX 18-JUL-1986; 86US-00886260.
PR
XX
XX (UPJO ) UPJOHN CO.
PA
XX (PETR/) PETROVSKIS E A.
PA
XX
XX Post LE, Petrovskis EA, Timmins JG;
PI
XX
XX WPI; 1987-108699/15.
DR
XX N-PSDB; AAN70994.
DR
XX
XX Pseudo-rabies virus protein - produced from recombinant DNA and used to
PT produce vaccine and detect animals infected with virulent virus.
PT
XX
XX Claim 8; Page 59; 67pp; English.
PS
XX
XX The PRV glycoprotein product may be used as in vaccination of animals
CC such as swine, sheep and goats against infection by the virus. The
CC protein may be produced from a transformed expression system such as
CC E.coli, yeast or CHO cells operatively linked to a suitable expression
CC control sequence. Glycoproteins gpi and gp63 may be used to distinguish
CC between infected and vaccinated animals. See also AAN70993-5
CC
XX
XX Sequence 350 AA;
SQ
Query Match 7.7%; Score 143.5; DB 1; Length 350;
Best Local Similarity 22.8%; Pred. No. 4.3e-05;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;
QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQRPKP 65
DB 15 LLLAALTTLAALTFRVGVLFPGAGSVHVAGSAVLVPGDAPNLTDGTLLEGPSPN- 73
QY 66 YSGTVRVAFRSDITNQCYQLSEERFENCNTHRSSVFGC--KVTEYTFAS-----NRL 118
DB 74 YSGRVEL-LRLDPKRCYTRYEAAEYDLCPRVHHEAFRGCLRKREPLARRASAAVEARL 132
QY 119 TGPFPFKLIRPNRPNDSGMFYIVRLDDTKFIDVFAIQLSVYQFANTATRGL---- 174
DB 133 -----LFVSRPAPPDAGSYVLRVNGT---TDLFVLTLV-----PPRGRPHP 174
QY 175 --YSKASCRFTGLPTVQLEAYLRTEESWRNQAQV--ATEATTSA-----EATPTTPV 224
DB 175 TPSSADECR----FVV-----GSHDSLAVVDPDAEDAVTTTPIEPDPTTPAPP 221
QY 225 TATSAS-----ELEAHTFPWLENGVDHYEPTPANENSVTVRLGWNPSFLIGTVTA 277
DB 222 RGTGATPEPRSDDEEDE-----EGATTAMPVPGTLDANGTM---VLNASVWSRVLL 271
QY 278 AVVSATIG-----LVIVISIVRNWCTPHRKLDTVQSDDERSQTRRESK 323
DB 272 AANATAGARGPGKIAMVLGPTIIVLLIFLGGVACAAARCARGIASTGRDPGAARSTRR 331

```

RESULT 12  
 AAR63143  
 ID AAR63143 standard; protein; 350 AA.  
 XX  
 AC AAR63143;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 12-JUN-1995 (first entry)  
 XX  
 XX Glycoprotein 63 (gp63) of pseudorabies virus.  
 DE  
 XX pseudorabies virus; PRV; glycoprotein; gp50; gI; gp63; determination;  
 KW detection; vaccine; infected animal; isolation; cloning; virulence.  
 KW  
 XX Pseudorabies virus.  
 OS  
 XX US5352575-A.  
 PN  
 XX 04-OCT-1994.  
 PD  
 XX 20-APR-1990; 90US-00513282.  
 PF  
 XX 04-OCT-1985; 85US-00784787.  
 PR  
 XX 26-NOV-1985; 85US-00801799.  
 PR  
 XX 26-MAR-1986; 86US-00844113.  
 PR  
 XX 16-JUL-1986; 86US-00886260.  
 PR  
 XX 29-JUN-1987; 87US-00100817.  
 PR  
 XX (UPJO ) UPJOHN CO.  
 PA  
 XX Timmins JG, Post LE, Petrovskis EA;  
 PI  
 XX WPI; 1994-316176/39.  
 DR  
 XX N-PSDB; AAR73489.  
 DR  
 XX  
 XX Identifying animals vaccinated against pseudorabies virus - by detecting  
 PT the absence of gI or gp63 antibodies in serum to distinguish vaccinated  
 PT from infected animals.  
 PT  
 XX Example 3; Col 21-24; 21pp; English.  
 PS  
 XX AAR63143 shows the protein sequence of gp63 (glycoprotein) encoded by  
 CC AAR73489, isolated from Pseudorabies virus (PRV). The DNA and protein  
 CC sequences of the invention are useful in a method for distinguishing an  
 CC animal vaccinated with a PRV vaccine lacking glycoprotein gI, gp63 or  
 CC gp50 from an animal infected with a virulent wild-type PRV without  
 CC sacrificing the animal. The method is used to test animals such as swine,  
 CC cattle, sheep and goats. (see AAR63142 and AAR63144). (Updated on 25-MAR-  
 CC 2003 to correct PF field.)  
 CC  
 XX Sequence 350 AA;  
 SQ

Query Match 7.7%; Score 143.5; DB 2; Length 350;  
 Best Local Similarity 22.8%; Pred. No. 4.3e-05;  
 Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;  
 QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQRPKP 65  
 DB 15 LLLAALTTLAALTFRVGVLFPGAGSVHVAGSAVLVPGDAPNLTDGTLLEGPSPN- 73  
 QY 66 YSGTVRVAFRSDITNQCYQLSEERFENCNTHRSSVFGC--KVTEYTFAS-----NRL 118  
 DB 74 YSGRVEL-LRLDPKRCYTRYEAAEYDLCPRVHHEAFRGCLRKREPLARRASAAVEARL 132  
 QY 119 TGPFPFKLIRPNRPNDSGMFYIVRLDDTKFIDVFAIQLSVYQFANTATRGL---- 174  
 DB 133 -----LFVSRPAPPDAGSYVLRVNGT---TDLFVLTLV-----PPRGRPHP 174  
 QY 175 --YSKASCRFTGLPTVQLEAYLRTEESWRNQAQV--ATEATTSA-----EATPTTPV 224  
 DB 175 TPSSADECR----FVV-----GSHDSLAVVDPDAEDAVTTTPIEPDPTTPAPP 221

QY 225 TATSAS-----PLEAEHFTPPHLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVVA 277  
DB 222 RGTGATPEPRSDDEEED-----EGATTAMTPVPGTLDANGTM---VLNASVSVRLL 271  
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDTVSDDEERSQTRRESRK 323  
DB 272 AAANATAGARGPGKIAMVGLPTIVLLIFLGVCACARRCARGIASTGRDPGARRSTR 331

RESULT 13  
AAE05396  
ID AAE05396 standard; protein; 350 AA.  
AC AAE05396;  
XX  
DT 24-SEP-2001 (first entry)  
XX  
DE Pseudorabies virus (PRV) glycoprotein gp63.  
XX  
KW Pseudorabies virus; PRV; glycoprotein; gp63; immunostimulant; vaccine;  
KW PRV infection.  
XX  
OS Pseudorabies virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 131 /note= "Encoded by GCG"  
FT Misc-difference 132 /note= "Encoded by GCG"  
FT Misc-difference 144 /note= "Encoded by GCG"  
FT Misc-difference 207 /note= "Encoded by GCG"  
FT Misc-difference 222 /note= "Encoded by GCG"  
FT Misc-difference 223 /note= "Encoded by GCG"  
FT Misc-difference 284 /note= "Encoded by GCG"  
FT Misc-difference 334 /note= "Encoded by GCG"  
FT /note= "Encoded by GCC"  
XX  
PN US6261563-B1.  
XX  
PD 17-JUL-2001.  
XX  
PF 07-JUN-1995; 95US-00485287.  
XX  
PR 26-NOV-1985; 85US-00801799.  
PR 26-MAR-1986; 86US-00844113.  
PR 16-JUL-1986; 86US-00886260.  
PR 28-AUG-1986; 86WO-US001761.  
PR 29-JUN-1987; 87US-00100817.  
PR 20-APR-1990; 90US-00513282.  
PR 21-JUN-1994; 94US-00262813.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Petrovskis EA, Post LR, Timmins JG;  
XX  
DR WPI; 2001-450478/48.  
DR N-PSDB; AAD10195.  
XX  
PT Polypeptide comprising a pure and isolated pseudorabies virus gp63  
PT polypeptide or its fragments, useful as a vaccine for protecting animals  
PT against pseudorabies virus infection.  
XX  
PS Claim 1; Col 39-40; 21pp; English.  
XX  
CC The invention relates to Pseudorabies virus (PRV) glycoproteins gp50,  
CC gp63, gI and their corresponding DNA molecules. These glycoproteins are  
CC used as vaccines for protecting animals against PRV infection. The  
CC invention also relates to methods for protecting animals against PRV

CC infections and methods for distinguishing between infected and vaccinated  
CC animals. The present sequence is Pseudorabies virus (PRV) glycoprotein  
CC gp63  
XX  
SQ Sequence 350 AA;  
Query Match 7.7%; Score 143.5; DB 4; Length 350;  
Best Local Similarity 22.8%; Fred. No. 4.3e-05;  
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;  
QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQLOLFPMPGQRPKHP 65  
DB 15 LLLAALTTLAALTTPRVGGVIFRGAGVSVHVAGSAILVPGDAPNLITDGLLFLGSPSN- 73  
QY 66 YSGTVRVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFSSA-----NRL 118  
DB 74 YSRVEL-LRLDPKACYTREYAAEYDLCPRVHEAPRGCLRKREPLARRASAAVEARRL 132  
QY 119 TGPPHPFKLTRNRPNDNGMFYVIVLDDTKBIDVFAIQLSVYQFANTATRGL----- 174  
DB 133 -----LFVSRPAPPDAGSVLVRVNGT---TDLFVLTVLV-----PPRGRPHHP 174  
QY 175 --YSKASCTFGLPTVQLEAVLRTEESWRNQAVV--ATEATTTS-----AEATTPTPV 224  
DB 175 TPSSADECR----PVV-----GSHDSLVRVVDPAEDAVFTTPPPIEPPTTPAPP 221  
QY 225 TATSAS-----ELEAEHFTPPHLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVVA 277  
DB 222 RGTGATPEPRSDDEEED-----EGATTAMTPVPGTLDANGTM---VLNASVSVRLL 271  
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDTVSDDEERSQTRRESRK 323  
DB 272 AAANATAGARGPGKIAMVGLPTIVLLIFLGVCACARRCARGIASTGRDPGARRSTR 331

RESULT 14  
AAR80637  
ID AAR80637 standard; protein; 317 AA.  
XX  
AC AAR80637;  
XX  
DT 21-DEC-1995 (first entry)  
XX  
DE Bovine herpes virus-1 mutant ORF1 product.  
XX  
KW BHV-1; vaccine; gIV gene; antigen.  
XX  
OS Bovine herpesvirus.  
XX  
PN EP663403-A1.  
XX  
PD 19-JUL-1995.  
XX  
PF 18-NOV-1994; 94EP-00203361.  
XX  
PR 23-NOV-1993; 93EP-00203274.  
XX  
PA (ALKU ) AKZO NOBEL NV.  
XX  
PI Keil G;  
XX  
DR WPI; 1995-247499/33.  
DR N-PSDB; AAQ99001.  
XX  
PT New bovine herpes virus with mutation in ORF1 and polyA signal for gIV  
PT gene - useful as a vaccine, esp. where the mutation is insertion of  
PT sequence encoding an heterologous antigen.  
XX  
PS Claim 2; Page 13-14; 28pp; English.  
XX  
CC To produce a new BHV-1 mutant, DNA encoding gIV (flanked on both sides by  
CC a sequence that allows homologous recombination) and a genotypically-neg.  
CC BHV-1 (or its DNA) is introduced into a host cell. Since the gIV-neg.



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:39:45 ; Search time 30 Seconds  
(without alignments)  
622.953 Million cell updates/sec

Title: US-09-993-777-7  
Perfect score: 1863  
Sequence: 1 MASLIGTALLAATLAPFGA.....VELVAIVNPNSALSPDSIKM 362

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A-COMB.pep:\*

2: /cgn2\_6/prodata/2/iaa/5B-COMB.pep:\*

3: /cgn2\_6/prodata/2/iaa/6A-COMB.pep:\*

4: /cgn2\_6/prodata/2/iaa/6B-COMB.pep:\*

5: /cgn2\_6/prodata/2/iaa/PCTUS-COMB.pep:\*

6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1863	100.0	362	5	PCT-US96-03916-7
2	1863	100.0	362	5	PCT-US96-03916-68
3	186	8.4	384	4	US-08-504-617-2
4	152	8.2	355	4	US-08-709-731A-28
5	148	7.9	317	2	US-08-344-833-2
6	148	7.9	364	1	US-08-680-726A-56
7	148	7.9	364	3	US-09-092-409-56
8	145	7.8	370	3	US-08-911-321-10
9	143.5	7.7	350	6	5352575-7
10	140.5	7.5	380	3	US-08-924-345-2
11	112	6.0	387	4	US-09-175-928-2
12	108	5.8	713	1	US-08-190-802A-63
13	108	5.8	713	3	US-08-477-346-63
14	108	5.8	713	4	US-08-473-089-63
15	108	5.8	713	4	US-08-487-072A-63
16	108	5.8	713	4	US-09-108-857-3
17	107.5	5.8	100	3	US-09-047-125-27
18	107.5	5.8	100	3	US-07-736-335B-27
19	103.5	5.6	94	1	US-08-150-203A-6
20	103.5	5.6	94	1	US-08-454-730-6
21	103.5	5.6	94	4	US-08-949-788-6
22	99	5.3	97	2	US-09-047-125-25
23	99	5.3	97	3	US-07-736-335B-25
24	99	5.3	820	4	US-09-489-039A-10142
25	98	5.3	503	1	US-07-946-497-2
26	98	5.3	503	1	US-08-483-322-2
27	98	5.3	503	2	US-08-478-882-2

28 97 5.2 1220 2 US-08-843-530B-36 Sequence 36, Appli

29 94.5 5.1 307 4 US-09-197-970B-3 Sequence 3, Appli

30 93.5 5.0 878 4 US-09-556-706B-2 Sequence 2, Appli

31 93.5 5.0 907 3 US-08-783-774-2 Sequence 2, Appli

32 93.5 5.0 907 4 US-09-328-599A-1 Sequence 1, Appli

33 93.5 5.0 907 5 PCT-US95-04611A-19 Sequence 19, Appli

34 92.5 5.0 547 1 US-08-314-615-1 Sequence 1, Appli

35 92.5 5.0 547 1 US-08-314-362-1 Sequence 1, Appli

36 92.5 5.0 547 1 US-08-433-010-1 Sequence 1, Appli

37 92.5 5.0 547 1 US-08-482-882-1 Sequence 1, Appli

38 92.5 5.0 547 2 US-08-483-389-1 Sequence 1, Appli

39 92.5 5.0 547 2 US-08-487-113D-1 Sequence 1, Appli

40 92.5 5.0 547 2 US-08-473-503-1 Sequence 1, Appli

41 92.5 5.0 547 2 US-08-483-932-1 Sequence 1, Appli

42 92.5 5.0 547 2 US-08-720-420A-1 Sequence 1, Appli

43 92.5 5.0 547 3 US-08-714-017-1 Sequence 1, Appli

44 92.5 5.0 547 3 US-08-863-790-1 Sequence 1, Appli

45 92.5 5.0 547 3 US-08-475-680-1 Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
PCT-US96-03916-7  
; Sequence 7, Application PC/TUS9603916  
; GENERAL INFORMATION:  
; APPLICANT: Wild, Martha A.  
; APPLICANT: Cochran, Mark D.  
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/03916  
; FILING DATE: 23-MAR-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,597  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 39116-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-03916-7

Query Match 100.0%; Score 1863; DB 5; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.3e-188;  
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVAPRPETATIQQLFFMFGQ 60  
DB 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVAPRPETATIQQLFFMFGQ 60

QY 61 RPKPXSQTVRVAFRSDITNQCYOELSEERFENCITHRSSSVFVCKVTEYTFSSASNLRTG 120  
DB 61 RPKPXSQTVRVAFRSDITNQCYOELSEERFENCITHRSSSVFVCKVTEYTFSSASNLRTG 120  
QY 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPEIDVFAIQLSVYQFANTAATRGLYSKASC 180  
DB 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPEIDVFAIQLSVYQFANTAATRGLYSKASC 180  
QY 181 RTGLPTVQLVQLAYLRTEESWRNWOAYVATEATTTSABATTPPTVATASASELEAEHTFP 240  
DB 181 RTGLPTVQLVQLAYLRTEESWRNWOAYVATEATTTSABATTPPTVATASASELEAEHTFP 240  
QY 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNCTP 300  
DB 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNCTP 300  
QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360  
DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360  
QY 361 KM 362  
DB 361 KM 362

## RESULT 2

PCT-US96-03916-68  
; Sequence 68, Application PC/TUS9603916  
; GENERAL INFORMATION:  
; APPLICANT: Wild, Martha A.  
; APPLICANT: Cochran, Mark D.  
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/03916  
; FILING DATE: 23-MAR-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,597  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 39116-A  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-03916-68

Query Match 100.0%; Score 1863; DB 5; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.3e-188;  
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHHIVIVAPRPEATIQQLFFPMGQ 60  
DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHHIVIVAPRPEATIQQLFFPMGQ 60  
QY 61 RPKPXSQTVRVAFRSDITNQCYOELSEERFENCITHRSSSVFVCKVTEYTFSSASNLRTG 120  
DB 61 RPKPXSQTVRVAFRSDITNQCYOELSEERFENCITHRSSSVFVCKVTEYTFSSASNLRTG 120  
QY 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPEIDVFAIQLSVYQFANTAATRGLYSKASC 180  
DB 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPEIDVFAIQLSVYQFANTAATRGLYSKASC 180  
QY 181 RTGLPTVQLVQLAYLRTEESWRNWOAYVATEATTTSABATTPPTVATASASELEAEHTFP 240  
DB 181 RTGLPTVQLVQLAYLRTEESWRNWOAYVATEATTTSABATTPPTVATASASELEAEHTFP 240  
QY 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNCTP 300  
DB 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNCTP 300  
QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360  
DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360  
QY 361 KM 362  
DB 361 KM 362

## RESULT 3

US-08-504-617-2  
; Sequence 2, Application US/08504617  
; Patent No. 6521236  
; GENERAL INFORMATION:  
; APPLICANT: Willemse, Martha Jacoba  
; APPLICANT: Sondermeijer, Paulus Jacobus Antonius  
; TITLE OF INVENTION: Vector vaccines of recombinant  
; TITLE OF INVENTION: Feline herpesviruses  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 6521236el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: U.S.A.  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/504,617  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/211,150  
; FILING DATE: March 22, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELEPHONE: (301) 258-5200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 384 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; OTHER INFORMATION: /label= ORF-1  
US-08-504-617-2





TELEPHONE: (301) 258-5200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 317 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-834-833-2

Query Match 7.9%; Score 148; DB 2; Length 317;  
Best Local Similarity 27.1%; Pred. No. 4.1e-07;  
Matches 68; Conservative 34; Mismatches 125; Indels 24; Gaps 11;

QY 1 MASLLGTLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQLFFMP 58  
DB 1 MERLLLMVLAARAAP--ARSLVYRGEAVGLRADGPVAVFVHP-TDATALRGLRIFLE 57  
QY 59 GQRP-HKPSGTVRVAFRSDITNQCQELSEERFENCNTHRSSSVFVGC--KVTEYTFASNLTG 117  
DB 58 HOLPAGRRYNGTVEL-LRYHAAGDCFVMLQTTAFASCPRVANDAFRSLHADTRPARSER 116  
QY 118 LTGP--PHPFKLTIRPNRNDSGMPYIVRL---DDTKEPIDVFAIQLSVYQFAN---- 167  
DB 117 RASAAVENHVLFSIARPIIDSLGLYFLRVGYGTAGSERRRDVFPPLAAVHSGFGGDP 176  
QY 168 --TAATRGYS--KASCTFGLPTVQLEAYLRTESWRNWOAYVATEATTTSAAETPTP 223  
DB 177 EAAAHAPGTVAEVACRER-GLDASSASLYDRAALAAPGAA--TTPGPTASSESQAATP 233  
QY 224 VTATSASLEA 234  
DB 234 ERVDETTEVEA 244

## RESULT 6

US-08-680-726A-56  
; Sequence 56, Application US/08680726A  
; Patent No. 5804197  
; GENERAL INFORMATION:  
; APPLICANT: Haanes, Elizabeth J.  
; APPLICANT: Frank, Rexann S.  
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES  
; NUMBER OF SEQUENCES: 92  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 12-JUL-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-46-C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-680-726A-56

Query Match 7.9%; Score 148; DB 1; Length 364;  
Best Local Similarity 27.3%; Pred. No. 5.1e-07;  
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLAALATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQLFFMPGQR-P 62  
DB 16 LLITMFLPILFLYGVNGFVYKGTIYSMTINTSSGFSIFPDDKFIIVSGRLFLDDQHLS 75  
QY 63 HKPSGTVRVAFRSDITNQCQELSEERFENCNTHRSSSVFVGC--KVTEYTFASNLTG 120  
DB 76 VNNYSGTIEFI---HFNNSCYTVYQTIETFPFNNAPRSLCKVKVSKHESQLRINS 132  
QY 121 PHPFKLTIRPNRNDSGMPYIVRLDDTKEPIDVFAIQLSVYQF 165  
DB 133 LENGVLLEITNPKNDSGVFIIVQLENNK--TDVFGIPAPIYSP 175

## RESULT 7

US-09-092-409-56  
; Sequence 56, Application US/09092409  
; Patent No. 6159478  
; GENERAL INFORMATION:  
; APPLICANT: Haanes, Elizabeth J.  
; APPLICANT: Frank, Rexann S.  
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES  
; NUMBER OF SEQUENCES: 92  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/09/092,409  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/680,726  
; FILING DATE: 12-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-46-C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-092-409-56

Query Match 7.9%; Score 148; DB 3; Length 364;  
Best Local Similarity 27.3%; Pred. No. 5.1e-07;  
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLAALATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQLFFMPGQR-P 62  
DB 16 LLITMFLPILFLYGVNGFVYKGTIYSMTINTSSGFSIFPDDKFIIVSGRLFLDDQHLS 75  
QY 63 HKPSGTVRVAFRSDITNQCQELSEERFENCNTHRSSSVFVGC--KVTEYTFASNLTG 120  
DB 76 VNNYSGTIEFI---HFNNSCYTVYQTIETFPFNNAPRSLCKVKVSKHESQLRINS 132

QY 121 PPHPKLTIRNPRNDGMPYVIVRLDDTKPIIDVFAIQLSVYQF 165  
Db 133 IENGVLLEITNPXPNDSGVYFIRVQLENNK--TDVFGIPAFIYSF 175

US-08-911-321-10  
Sequence 10, Application US/08911321  
Patent No. 6010703  
GENERAL INFORMATION:  
APPLICANT: Roger K. Maes and Stephen J. Spatz  
TITLE OF INVENTION: Recombinant Poxvirus  
TITLE OF INVENTION: Vaccine Against  
TITLE OF INVENTION: Feline Rhinotracheitis  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/911,321  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/096,183  
FILING DATE: July 26, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: MSU 4.1-166  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 6010703e  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE:  
DESCRIPTION: Polypeptide  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: Feline herpesvirus-1  
STRAIN: 1  
INDIVIDUAL ISOLATE: C-27  
CELL TYPE: N/A  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD: Deduced Sequence  
OTHER INFORMATION: 91

US-08-911-321-10  
Query Match 7.8%; Score 145; DB 3; Length 370;  
Best Local Similarity 22.0%; Pred. No. 1.1e-06;  
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;  
QY 5 LGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRE-ATLQQLFPMPGQ-RP 62  
Db 1 MSSIAFVILMAIGTVIGYVRGDHVSILHVDTSSTSGFVIYPTLENFTIYGHILFDLDDQPLP 60

QY 63 HKPYSGTVVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFASNRLTG 120  
Db 61 VNYNTGTLLEI--IHYNHSSCKYIVQVIEYSSCPVRNNAFRSCHUKISMEOYDQLSINTS 119  
QY 121 PPHPKLTIRNPRNDGMPYVIVRLDDTKPIIDVFAIQLSVYQFANTAATRGYSKASC 180  
Db 120 VETGMLLTITSPKMGDGIYALRVFNHNK-ADVFGLSVFVYSF---DTRGHRHADE 174  
QY 181 RTFG--LPTVO--LEAYLRTEESWRNWOAY---VATEATTTSAAEATTPPTVATATASELEA 234  
Db 175 NLNGEILLTPSPMETVYKV---NTPYDHMTVTTQTTSNKSMESPSNTSISC-----223  
QY 235 EHFTFPWLENGVDHYEPTPANENSNTVRL-----GTMSPTLIGVTVA 278  
Db 224 -----HTFQNDPNEGETLYTHLLNIAGNITYDDMVMVDGTTLKPRLL-----264  
QY 279 VVSATIGLIVIVISIVTRNNCTPHEKLDIT 306  
Db 265 ----DMGLNLSVTSSFKN--GNHAKMDT 286

RESULT 9  
5352575-7  
; Patent No. 5352575  
; APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.  
; TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/513,282  
; FILING DATE: 20-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 100,817  
; FILING DATE: 29-JUN-1987  
; APPLICATION NUMBER: 886,260  
; FILING DATE: 16-JUL-1986  
; APPLICATION NUMBER: 784,787  
; FILING DATE: 04-OCT-1985  
; APPLICATION NUMBER: 801,799  
; FILING DATE: 26-NOV-1985  
; APPLICATION NUMBER: 844,113  
; FILING DATE: 26-MAR-1986  
; SEQ ID NO: 7;  
; LENGTH: 350  
5352575-7

Query Match 7.7%; Score 143.5; DB 6; Length 350;  
Best Local Similarity 22.8%; Pred. No. 1.4e-06;  
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;  
QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPREATLQQLFPMPGQRPKP 65  
Db 15 LLLAALTTLAALTPTRVGGVLFRTGAAVSVHVAGSALVPGDAPNLTIDGTLFLLEGSPSN- 73  
QY 66 YSGTVVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFAS-----NRL 118  
Db 74 YSGRVEL--LRLDPKACTREYAAEYDLCPRVHAEFGCLRKREPLARRASAAVEARRL 132  
QY 119 TGPFPFKLTIRNPRNDGMPYVIVRLDDTKPIIDVFAIQLSVYQFANTAATRGIL-----174  
Db 133 -----LFVSRPAPPDAGSVLRVRVNGT---TDLFVLTAIV-----PPRGRPHHP 174  
QY 175 --YSKASCTFGLPTVQLEAYLRTEESWRNWOAYV--ATEATTTS-----AEATTPTPV 224  
Db 175 TPSSADECR-----PVV-----GSWHDLSRVVDPAEDAVFTTTPPIEPPEPTTAPP 221  
QY 225 TATSAS-----ELEAEHFTFPWLENGVDHYEPTPANENSNTVRLGTMSTLIGVTVA 277  
Db 222 RTGATPEPRSRDEEED-----EGATTAMTPVPGTLDANGTM---VLNASVSVRL 271  
QY 278 AVYSATIG-----LVIVISIVTRNMCTPHRKLDIVSQDDEERSQTRRSRK 323  
Db 272 AANATAGAGPGKTAWLGTPIVILLFLGGVACAARCARCARGIASTGRDPCGAARRSTR 331

RESULT 10  
US-08-924-345-2  
; Sequence 2, Application US/08924345  
; Patent No. 6224878  
; GENERAL INFORMATION:  
; APPLICANT: LEUNG-PACK Patricia  
; APPLICANT: LEGASTLOIS Isabelle, Christine, Marie-Andree  
; APPLICANT: AUDONNET Jean-Christophe, Francis  
; APPLICANT: RIVIERE Michel, Emile, Albert  
; TITLE OF INVENTION: Mutants and vaccines of the Infectious  
; TITLE OF INVENTION: Bovine Rhinotracheitis virus  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LARSON AND TAYLOR  
; STREET: 727 SOUTH TWENTY-THIRD STREET  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/924,345  
; FILING DATE: 04-SEP-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/199,172  
; FILING DATE: 09-AUG-1994  
; APPLICATION NUMBER: FR 92 07930  
; FILING DATE: 26-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SARRO, THOMAS P.  
; REGISTRATION NUMBER: 19396  
; REFERENCE/DOCKET NUMBER: XI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 920-7200  
; TELEFAX: (703) 892-8428  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-924-345-2

Query Match 7.5%; Score 140.5; DB 3; Length 380;  
Best Local Similarity 25.3%; Pred. No. 3.4e-06;  
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;  
QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPETATLQ--QLFFMP 58  
DB 1 MRCULLMWVLAARAAP--ARSLVYRGAVGLRAGDGVAFVHP-TDATALGRLLIFLE 57  
QY 59 GQRP-HKPYSGTVRVAFRSDITNOCYQELSEERFENCCTHRSSTSSVFGCKVTEYTPFSASNR 117  
DB 58 HQLPAGRYNGTVEL-LRYHAAGCFVMLQTTAFASCPRVANNAFSLHADTPARSR 116  
QY 118 LTGP--PHFPKLTIRNPNDSGMFYVIVRL-----DOTKEPIDVFAIQLSVYQFA----- 166  
DB 117 RASAAVENHLVFSIARPPIDSLGLYFLRVGYGTAGSERRRDVFLAAFFVHSGEGDP 176  
QY 167 -----NRAATGLYSKASCTFFGL-----PTVQLEAYLRTBSWR 201  
DB 177 EAAATPAPRSQRPAASGLTSSASLYDRALARSFQAPPPRPAARAGPRPRXVDE 236  
QY 202 NWQAVATEATTSAETTPPVATTSAS 230  
DB 237 TTEVEAATRAGSAFALTTPAGPTASPA 265

RESULT 11  
US-09-175-928-2  
; Sequence 2, Application US/09175928A  
; Patent No. 6312921  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallie, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Mi, Sha  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: 6006B.AJ172A  
; CURRENT APPLICATION NUMBER: US/09/175,928A  
; CURRENT FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 387  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-175-928-2  
Query Match 6.0%; Score 112; DB 4; Length 387;  
Best Local Similarity 19.2%; Pred. No. 0.0036;  
Matches 77; Conservative 58; Mismatches 129; Indels 138; Gaps 20;  
QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPETATLQQLFFMPGQ 60  
DB 12 LSLAGQSVVQVTI-PDGFVNVTVGSNVTLICI---YTTTVASREQLSIQWSFF----- 62  
QY 61 RPKPYSGTVRVAFRSDITNOCYQELSEERFENCCTHRSSTSSVFGCKVTEYTPFSASNRITG 120  
DB 63 --HKMEPISYF-----SQGQVAIG-----QFKDRIIG 92  
QY 121 PPHP--FKLTIRNPNDSGMFYVIVRLDDTKPEIDVFAIQLSVYQVFANTATRLGLYSKA 178  
DB 93 SNDPQNASITISHMQPADSGIY-----ICDWNPPDFLGQNGILNVSLVKP-----SKP 143  
QY 179 SCRTFGLP-----TVOLRAYLRTESWRNQAVATEATTSAEATTPPTVATTSASELEA 234  
DB 144 LCSVQGRPETHGTISLSCL-----SALGTPSPV----- 171  
QY 235 EHFTFPWLENGVDHYEPTPANENSNVTV-----RLG----- 265  
DB 172 ----YYW--HKLEGRDIVPKENFNPTTGILVIGLNTNFEQGYOCTAINRLGNSSCEID 225  
QY 266 -TMSPTLIGTVVAUVSATIGLVIVISIVTRNMCTPHKLDTVSODDEERSOTRESRKF 324  
DB 226 LTSSHPEVGIIVGALIGSLVGAIIISVV-----CFARNKAKAKAK--ERNSTIAE---L 276  
QY 325 GPMVACEINKGADQDS---ELVELVAIVNPSAL---SSPDSIK 361  
DB 277 EPMT--KINPRGSEAMPREDATQLEVLTPSIIHETGPDITQ 316  
RESULT 12  
US-08-190-802A-63  
; Sequence 63, Application US/08190802A  
; Patent No. 5519003  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: TUPl, Fig. 46  
US-08-190-802A-63

Query Match 5.8%; Score 108; DB 1; Length 713;  
Best Local Similarity 27.3%; Pred. No. 0.025;  
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRVAFRSDITNQCYQELSEERFENCSTRSSVFGCKVTEYTFSS---AS 115  
DB 86 QRDHQIASLTVCQQQQQQQQVQHLQQQQ-QQLAASASVPVAQPPATTSATAPAA 144  
QY 116 NRLTGPPHPFKLTIRNPRNDSGMFMVIVRLDDTKPEI-----D 154  
DB 145 NTTGSPSAF--PVQASRPNLVGS-----QLPTTLTPVSSNAQQQLPQQOQQOQQ 197  
QY 155 VFAIQLSVYQFANTAATGLYSKASCTFGTLPTVOL-EAYLRTESWRNQAIV-----A 208  
DB 198 QPPQVSVAPLNTAING---SPTSKEITTLPSVKAPESTLKETEPENNNTSKINDTGA 254

QY 209 TEATTTTSABAT--TPTPTVATASAELEAEHFTFPWLENGVDHYEPTP 253  
DB 255 TTATTTTATETIKPEEDATPAS-LHQDHYLVFPYNQR-AHNSKPIP 299

RESULT 13  
US-08-477-346-63  
Sequence 63, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: TUPl, Fig. 46  
US-08-477-346-63

Query Match 5.8%; Score 108; DB 3; Length 713;  
Best Local Similarity 27.3%; Pred. No. 0.025;  
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRVAFRSDITNQCYQELSEERFENCSTRSSVFGCKVTEYTFSS---AS 115  
DB 86 QRDHQIASLTVCQQQQQQQQVQHLQQQQ-QQLAASASVPVAQPPATTSATAPAA 144  
QY 116 NRLTGPPHPFKLTIRNPRNDSGMFMVIVRLDDTKPEI-----D 154  
DB 145 NTTGSPSAF--PVQASRPNLVGS-----QLPTTLTPVSSNAQQQLPQQOQQOQQ 197  
QY 155 VFAIQLSVYQFANTAATGLYSKASCTFGTLPTVOL-EAYLRTESWRNQAIV-----A 208  
DB 198 QPPQVSVAPLNTAING---SPTSKEITTLPSVKAPESTLKETEPENNNTSKINDTGA 254

QY 209 TEATTTTSABAT--TPTPTVATASAELEAEHFTFPWLENGVDHYEPTP 253  
DB 255 TTATTTTATETIKPEEDATPAS-LHQDHYLVFPYNQR-AHNSKPIP 299

RESULT 14  
US-08-473-089-63  
Sequence 63, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089

Search completed: March 8, 2004, 02:46:02  
Job time : 32 secs

RESULT 15  
US-08-487-072A-63  
US-08-487-072A-63  
Sequence 63, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Rog, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/POCKET NUMBER: 2550-0035.20  
TELECOMMUNICATION INFORMATION:



Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

Db 1 MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

QY 61 RHPKPYSGTVAVRSDITNOCYQELSEERFENCCHRSSVFGCKVTEYTFASNRLTG 120

Db 61 RHPKPYSGTVAVRSDITNOCYQELSEERFENCCHRSSVFGCKVTEYTFASNRLTG 120

QY 121 PPFPFKLTIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVQFANTAATRGLYSKASC 180

Db 121 PPFPFKLTIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVQFANTAATRGLYSKASC 180

QY 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

Db 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

Db 241 WLENGVDHYEPTPANENSNVTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

Db 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

QY 361 KM 362

Db 361 KM 362

RESULT 2

US-09-994-064-7

; Sequence 7, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/468,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 362 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-994-064-7

Query Match 100.0%; Score 1863; DB 10; Length 362;

Best Local Similarity 100.0%; Pred. No. 1.7e-179;

Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

Db 1 MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

QY 61 RHPKPYSGTVAVRSDITNOCYQELSEERFENCCHRSSVFGCKVTEYTFASNRLTG 120

Db 61 RHPKPYSGTVAVRSDITNOCYQELSEERFENCCHRSSVFGCKVTEYTFASNRLTG 120

QY 121 PPFPFKLTIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVQFANTAATRGLYSKASC 180

Db 121 PPFPFKLTIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVQFANTAATRGLYSKASC 180

QY 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

Db 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

Db 241 WLENGVDHYEPTPANENSNVTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

Db 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

QY 361 KM 362

Db 361 KM 362

RESULT 3

US-09-994-064-68

; Sequence 68, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/468,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 68:

; SEQUENCE CHARACTERISTICS:



LENGTH: 362 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-994-064-68

Query Match 100.0%; Score 1863; DB 10; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.7e-179;  
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLIGTALLAATAPFGAMGIVITGNHVSARIDDDHIVAVRPEATIQQLFFMPGQ 60  
DB 1 MASLIGTALLAATAPFGAMGIVITGNHVSARIDDDHIVAVRPEATIQQLFFMPGQ 60  
QY 61 RHKPVSGTVRVAFRSDITNOCYQELSEERFENCCHRSSTVFCCKVTEFSASNELTG 120  
DB 61 RHKPVSGTVRVAFRSDITNOCYQELSEERFENCCHRSSTVFCCKVTEFSASNELTG 120  
QY 121 PHPPFKLIRNPRNDSGMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGYSKASC 180  
DB 121 PHPPFKLIRNPRNDSGMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGYSKASC 180  
QY 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPPVTAATSAEAEHFTFP 240  
DB 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPPVTAATSAEAEHFTFP 240  
QY 241 WLENGVDHYEPTPANENSNVTVLGTMSPTLIGTVAAVGSATIGLVIVISIVTRNCTP 300  
DB 241 WLENGVDHYEPTPANENSNVTVLGTMSPTLIGTVAAVGSATIGLVIVISIVTRNCTP 300  
QY 301 HRKLTVSQDDEERSQTRRSKFGPMVACINKGADQDSELVELVAIVNPSALSPPDSI 360  
DB 301 HRKLTVSQDDEERSQTRRSKFGPMVACINKGADQDSELVELVAIVNPSALSPPDSI 360  
QY 361 KM 362  
DB 361 KM 362

RESULT 4  
US-10-156-275-56  
Sequence 56, Application US/10156275  
Publication No. US20030049844A1  
GENERAL INFORMATION:  
APPLICANT: Haanes, Elizabeth J.  
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES  
NUMBER OF SEQUENCES: 92  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/156,275  
FILING DATE: 28-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/399,118  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/680,726  
FILING DATE: 12-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-46-C1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 364 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 56:  
US-10-156-275-56

Query Match 7.9%; Score 148; DB 14; Length 364;  
Best Local Similarity 27.3%; Pred. No. 4.9e-06;  
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LIGTLALLAATAPFGAMGIVITGNHVSARIDDDHIVAVRPEATIQQLFFMPGQ-P 62  
DB 16 LLITMFLFLEFLYGVNGFVKQYISMLNTSSGFSIFPDDKFIYSGRLFLDDQHL 75  
QY 63 HKPVSGTVRVAFRSDITNOCYQELSEERFENCCHRSSTVFCCKVTEFSASNELTG 120  
DB 76 VNNYSGTIEFT--HFNNSCYTVVQTIYFSCPRIFNNAFRSLKVKYKHESQLRINS 132  
QY 121 PHPPFKLIRNPRNDSGMFYVIRLDDTKPIDVFAIQLSVYQF 165  
DB 133 IENGVLLEITNPRNDSGVIFIRVQLENNK--TDVGIAPAFIYGF 175

RESULT 5  
US-10-125-692-10  
Sequence 10, Application US/10125692  
Publication No. US2003004429A1  
GENERAL INFORMATION:  
APPLICANT: Aderem, Alan  
APPLICANT: Hayashi, Fumitaka  
APPLICANT: Smith, Kelly D.  
APPLICANT: Underhill, David M.  
APPLICANT: Ozinsky, Adrian  
TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods  
FILE REFERENCE: P-IS 5155  
CURRENT APPLICATION NUMBER: US/10/125,692  
CURRENT FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: US 60/285,477  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 612  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-125-692-10

Query Match 6.2%; Score 115; DB 14; Length 612;  
Best Local Similarity 20.6%; Pred. No. 0.024;  
Matches 76; Conservative 54; Mismatches 133; Indels 106; Gaps 16;

QY 19 GAGGIVITGNHVSARIDDDHIVAVRPEATIQQLF---FMPGQPHKPYSGTVRVAFR 75  
DB 93 GSFPFLII--NKLKEDSQTYICELENRKE-EVELWFKVTFSPGTSLLQGSJLTLD 149  
QY 76 SDITNOCYQELSEERFENCCHRSSTVFCCKVTEYFSAENRNLGTPHPFKLIRNPRN 135  
DB 150 SKVSN---PLTE-----CKHKGKGVSGSKV-----LSMSNLRVQ 181  
QY 136 DSGMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGYSKASCRTFGLPTVQLEAYLR 195  
DB 182 DSDFWNCTVILDQKK---NWFGMLSVLGFQSTAIT-AKSEGESAEFFPLNFAE---- 233  
QY 196 TEESWRN--WQAYVATEATTTSAEATTPPVTAATSAEAEHFTFPWLENGVDHYEPT 253  
DB 234 -ENCGELMWKA-----EKDSFFQFQWISFSIKKKEVS 265





Db 198 LSLTPSLP--BEATGLTTPESKEGPILTAESVTLPSDSWSVSTSDTAVLLTSKES 255  
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSLV 343  
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTTKTGQMDG--IPMSKMNEMPIS 313  
Qy 344 ELVAIVNPS 352  
Db 314 QLLMIAPS 322

## RESULT 11

US-10-188-012-33  
; Sequence 33, Application US/10188012  
; Publication No. US20030124114A1  
; GENERAL INFORMATION:  
; APPLICANT: McIntire, Jennifer Jones  
; APPLICANT: Umetsu, Dale T.  
; APPLICANT: Dekruyff, Rosemarie  
; APPLICANT: Kuchroo, Vijay  
; APPLICANT: Freeman, Gordon J.  
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
; FILE REFERENCE: STAN-235  
; CURRENT APPLICATION NUMBER: US/10/188,012  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(378)  
; OTHER INFORMATION: TIM-4, allele 1  
US-10-188-012-33

Query Match 5.5%; Score 102.5; DB 14; Length 378;  
Best Local Similarity 23.6%; Pred. No. 0.21;  
Matches 73; Conservative 36; Mismatches 133; Indels 67; Gaps 13;  
Qy 65 PYSCTVRVAFRSDITNQCQYQELSEPFENCTHRSSSVFVGCKVTEYTFPSASNLGTG--PP 122  
Db 60 PYSCKEALIRTD-----GMRVTSRK-SAKYRLOQTIPR 92  
Qy 123 HPFKLTIRNPRNDSGMFYVIVRL-----DDTKEPIDVFAIOLSVYQFANTAATRGLYSKA 178  
Db 93 GDSVLTILNPESDSGVYCCRIEVPGFNDVK-----INVRNL-QRASTTTHRTATTTT 146  
Qy 179 SCRTFGLPTVQLEAYLRTEESWRNMQAYVAEATTTSAEATTPP-----VTATSASE 231  
Db 147 RRTTTSPT-----TTRQMTTTPAALPTTVTT-PDLTTGTPLQMTTIAVFTTANTC 197  
Qy 232 LEAEHFTFPWLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVAAVAVSATIGLVVIS 291  
Db 198 LSLTPSLP--BEATGLTTPESKEGPILTAESVTLPSDSWSVSTSDTAVLLTSKES 255  
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSLV 343  
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTTKTGQMDG--IPMSKMNEMPIS 313  
Qy 344 ELVAIVNPS 352  
Db 314 QLLMIAPS 322

## RESULT 12

US-09-813-153-138  
; Sequence 138, Application US/09813153

; Publication No. US20030045459A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 67 Human secreted proteins  
; FILE REFERENCE: PZ023  
; CURRENT APPLICATION NUMBER: US/09/813,153  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US/09/363,044  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: 60/073,160  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,159  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,165  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,164  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,167  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,162  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,161  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,170  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 298  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 138  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (379)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-813-153-138

Query Match 5.5%; Score 102.5; DB 10; Length 379;  
Best Local Similarity 23.6%; Pred. No. 0.21;  
Matches 73; Conservative 36; Mismatches 133; Indels 67; Gaps 13;  
Qy 65 PYSCTVRVAFRSDITNQCQYQELSEPFENCTHRSSSVFVGCKVTEYTFPSASNLGTG--PP 122  
Db 60 PYSCKEALIRTD-----GMRVTSRK-SAKYRLOQTIPR 92  
Qy 123 HPFKLTIRNPRNDSGMFYVIVRL-----DDTKEPIDVFAIOLSVYQFANTAATRGLYSKA 178  
Db 93 GDSVLTILNPESDSGVYCCRIEVPGFNDVK-----INVRNL-QRASTTTHRTATTTT 146  
Qy 179 SCRTFGLPTVQLEAYLRTEESWRNMQAYVAEATTTSAEATTPP-----VTATSASE 231  
Db 147 RRTTTSPT-----TTRQMTTTPAALPTTVTT-PDLTTGTPLQMTTIAVFTTANTC 197  
Qy 232 LEAEHFTFPWLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVAAVAVSATIGLVVIS 291  
Db 198 LSLTPSLP--BEATGLTTPESKEGPILTAESVTLPSDSWSVSTSDTAVLLTSKES 255  
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSLV 343  
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTTKTGQMDG--IPMSKMNEMPIS 313  
Qy 344 ELVAIVNPS 352  
Db 314 QLLMIAPS 322

## RESULT 13

US-09-864-408A-5760  
; Sequence 5760, Application US/09864408A  
; Publication No. US2004000947A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shinkets, Richard A.

```

; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5760
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5760

Query Match      5.5%; Score 102; DB 11; Length 234;
Best Local Similarity 26.9%; Pred. No. 0.11;
Matches 61; Conservative 30; Mismatches 90; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRAFRSDI---TNQCYQELSEERFENCNTHRSSSVFGCKVTEYTFPS-----AS 115
Db 15 QRDQIASLTVQOQOQQOQQOQQOHLQOQO-QQLAASASVPVQAQPPATTSATATPAA 73
QY 116 NRLTGPHPFKLTIRNRPNDSGMVFYVIVRLDDTKEPIDVFAIQLSVYQFANTATG 173
Db 74 NITGSPSAP--PVQASRPNLVGS-----QLPTTLPWSSNAQQQLPQQOQQOQQO 126
QY 155 VPAQLSVQFANTATGELYSKASCTFGLPTVQL--EAYLTEESWENWQAVY-----A 208
Db 127 QPPQVQVAPLSNTAING---SPTSKEITLPSVKAPESTLKTETEPENNTSKINTGSA 183
QY 209 TEATTSAGAT--TPTVPTATSALEAEHFTFPWLENGVDHYEPTP 253
Db 184 TTAITTTATEIKPKEDATPAS--LQDHYLVFPYQNR--ANHSKPIP 228

RESULT 14
US-10-295-027-1308
; Sequence 1308, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08

```

```

; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1308
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1308

Query Match      5.4%; Score 101; DB 15; Length 413;
Best Local Similarity 19.7%; Pred. No. 0.34;
Matches 80; Conservative 61; Mismatches 144; Indels 122; Gaps 21;

QY 1 MASLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVAVRPEATIQQLFFMPGQ 60
Db 12 LSLAGQSVVQVTI--PDGFVNVTGVSNTLICI---YTTTVAEREQLSIQWSPFH---K 64
QY 61 RPKPYSGTVRAFRSDI---TNQCYQELSEERFENCNTHRSSSVFGCKVTEYTFPSAS 115
Db 65 KMEIPISSPWESEKWPDEAVKGLDGOQAEQLQIVF---SQGGQAVATG-----QPK 113
QY 116 NRLTGPHP--FKLTIRNRPNDSGMVFYVIVRLDDTKEPIDVFAIQLSVYQFANTATG 173
Db 114 DRITGNDPGNASITISHWQPADSGIY-----ICDVNPPDFLGQNGILNVSLVKP-- 166
QY 174 LYSKASCTFGLP-----TVQLEAYLTEESWENWQAVYATEATTSAGATPTPTATSA 229
Db 167 --SKPLCSVQGRPEGTHTISLCL-----SALGTPTSPV----- 197
QY 230 SELEAEHFTFPWLENGVDHYEPTPANENSVTV-----RLG-- 265
Db 198 -----YVW--HKLEGRDIVPKENFNPTTGLIVGLNLTNFEQGYVQCTAINRLGNS 246
QY 266 -----TMSPTLIGVTVAAVSATIGLVIVISIVTRNNCTPHRKLDTVSDDEERSQTRR 319
Db 247 SCEIDLTSSHPEVGIIVGALIGSLVGAIIISVW---CFARNKAKAKAK--ERNSKTIA 300
QY 320 ESRKFGPMVACEINKGAQDGS---ELVELVAIVNPSAL--SSPDSIK 361
Db 301 E---LEPMT--KINPRGSEAMPREDATQLEVTLPFSSIHETGPDTIQ 342

RESULT 15
US-10-282-122A-60126
; Sequence 60126, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60126
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-60126

Query Match      5.3%; Score 99; DB 12; Length 961;
Best Local Similarity 18.5%; Pred. No. 1.9;
Matches 48; Conservative 45; Mismatches 95; Indels 72; Gaps 9;

QY      20  AMGIVITGNHVSARIDDDHIVIVAPRPEATLQLFFMPGQRPKPYSGTVRVAFRSDIT 79
Db      427  ALDVVNIADQYDPOAIKORLAMTEQ-----NARIWISPOEPHNKTAYFYDAFYQVD-- 479

QY      80  NOCYOELSEBERPENTHRSSSVFVCKV-----TEYTFASNRLTGPPHP----- 124
Db      480  -----KISEQTPADWQHKSQAQLQLPALNPYIPDDFTLIKSDKAW--PHQQLILDBFTL 532

QY      125  -----FKLTIRNPNDSGMFYVIVIVELDDTKEPIDVFAIQLSVYQFAN 167
Db      533  RVVYAPSOYFASEPKADISLVRNPQAMDSARRQWFMALND-----YIAGIALDQLSN 585

QY      168  -----TAATRGLYSKASCRTFGLPTVQLE-----AYLRTEESW---RNWQAYVA 208
Db      586  QAAVGIGSFSTGANGNGLMVNANGYQHLPALFSDLQGYFSYTPTEEQLEQAKSWYACMM 645

QY      209  TEATTTSAEATTFPTVATTS 228
Db      646  DSAEKGYDQAIMPIQWYS 665
```

Search completed: March 17, 2004, 23:21:55  
Job time : 39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:31:35 ; Search time 42 Seconds  
(without alignments)  
829.079 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863  
Sequence: 1 MASLLGTLALLATLAPFGA.....VELVAIYNPSALSSPDSIKM 362

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	8.8	420	2 T42616	envelope protein -
2	162.5	8.7	354	1 VGBE67	glycoprotein D pre
3	155	8.3	353	1 C46113	glycoprotein D pre
4	153.5	8.2	424	1 VGBE89	glycoprotein gp63
5	143.5	7.7	350	1 VGBE83	glycoprotein gp63
6	140.5	7.5	380	2 S35785	glycoprotein I - b
7	125	6.7	356	2 JQ2352	glycoprotein I - t
8	124.5	6.7	683	2 T03146	probable glycoprot
9	115	6.2	457	1 RWM574	T-cell surface gly
10	113.5	6.1	390	1 QGBE77	glycoprotein I pre
11	111.5	6.0	149	2 A61162	glycoprotein I - e
12	108.5	5.8	355	2 C39725	hypothetical prote
13	108	5.8	713	2 JN0133	WD-40 repeat regul
14	107	5.7	372	1 QGBE88	glycoprotein I pre
15	99	5.3	372	2 F43674	US7 protein - huma
16	99	5.3	1036	2 S73601	protein p200 - Myc
17	98.5	5.3	1140	2 S73786	hypothetical prote
18	98	5.3	503	2 B38745	cell adhesion mole
19	98	5.3	507	2 S64507	probable membrane
20	97	5.2	1220	2 S48387	SIN1 protein - yea
21	96.5	5.2	814	1 A39752	fibroblast growth
22	95.5	5.1	797	1 VGBE41	glycoprotein X pre
23	95.5	5.1	867	2 T45463	membrane glycoprot
24	95	5.1	1777	2 T34369	hypothetical prote
25	94	5.0	796	2 T21460	hypothetical prote
26	93.5	5.0	645	2 T29818	hypothetical prote
27	93.5	5.0	907	1 QGBE21	membrane antigen g
28	93.5	5.0	3020	2 A43932	mucin 2 precursor,
29	93	5.0	1051	2 S55259	TIF1 protein - mou

30	92.5	5.0	1778	2 AF1116	internalin protein
31	92	4.9	1044	2 F95375	probable drug resi
32	91.5	4.9	510	2 T47374	hypothetical prote
33	91	4.9	365	2 A34424	CD44 membrane gly
34	91	4.9	836	2 S49940	cell division cont
35	90.5	4.9	547	1 S28904	intercellular adhe
36	90.5	4.9	785	2 T37329	probable oligopept
37	90.5	4.9	786	2 T19017	hypothetical prote
38	90.5	4.9	1429	2 T41699	C2-domain family p
39	90	4.8	233	2 C84193	hypothetical prote
40	90	4.8	798	2 T34248	hypothetical prote
41	89.5	4.8	826	2 G90283	hypothetical prote
42	89.5	4.8	866	2 T45462	membrane glycoprot
43	89.5	4.8	992	2 T38817	hypothetical prote
44	89	4.8	629	2 AE1525	probable peptidogl
45	89	4.8	678	2 S54308	DNA binding protei

ALIGNMENTS

RESULT 1  
T42616  
envelope protein - equine herpesvirus 4 (strain NS80S67)  
C:Species: equine herpesvirus 4  
A:Variety: strain NS80S67  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T42616  
R:Telford, B.A.; Watson, M.S.; Perry, J.; Cullinan, A.A.; Davison, A.J.  
J. Gen. Virol. 79, 1197-1203, 1998  
A:Title: The DNA sequence of equine herpesvirus-4.  
A:Reference number: Z22173; MUID:98264497; PMID:9603335  
A:Accession: T42616  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-420 <TSL>  
A:Cross-references: EMBL:AF030027; NID:G2605950; PIDN:AAC59593.1; PID:G2606031  
A:Experimental source: strain NS80S67  
C:Genetics:  
A:Note: 73  
C:Superfamily: pseudorabies virus glycoprotein gp63

Query Match	8.8%;	Score 164;	DB 2;	Length 420;
Best Local Similarity	25.5%;	Pred. No. 9.6e-06;		
Matches	96;	Conservative 50;	Mismatches 143;	Indels 88; Gaps 21;
QY	10	LLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQIQLQFEMEGOR-PHKPYS	67	
Db	13	LLAISMCL---ATAIIYRGHSMYLNASSEFAVIPKDKSLVVVGHMLFDGQRLPTNYS	69	
QY	68	GTVRVAFRSBITNQCQQLSEERFENCNTHRSSSVFVGC--KVTEYTFSSANRLTGPPHP	125	
Db	70	GLIEL-IHHVSRGCYSVIQTIYESCPRVANNAFSLHKTSHNHQDYFHVNTSVETNV	128	
QY	126	KLITRNPRNDGMFVIVRLDDTKPIDVFAIQLSVQF-ANTATRGLYSKASCRTPG	184	
Db	129	LLNITRPPQADGAYILRVKLNEA-PTADVGVSAFVYDLSQNTPEPVPETKPSNVET	187	
QY	185	LPTVQLAEYLRTESRNQAVYATZATTTSABAT---TPTPVATASASELEAEHETPW	241	
Db	188	RTAPAPANTSTK-----TGSNTSSOSTWLYTPTP-----RPA	221	
QY	242	LENGVDHYEETPANEN--SNVTVRL-----GTMSPT-----LIGVTVAASVATIG--	285	
Db	222	LET---HLTAPANEIVTWSGDTAMLCHGRFPSTAVPTIYHMLLGLT-----GNLPE	269	
QY	286	--LVIVISIVTRNMTFHRKLDIVS-----QDDEERSQTRRESR-KFGPMV-----ACEIN	333	
Db	270	DVLLIEDSEILR---TPPKPQTTSSRTGDDFKQTNSTSPKSRNKIVAMVVIPTACVLM	326	
QY	334	KGADQDSSELVELVAIYN	350	
Db	327	-----LLLVVVGAIIN	337	

```

RESULT 2
VCBE67
glycoprotein D precursor - human herpesvirus 3
N:Alternate names: glycoprotein IV
C:Species: human herpesvirus 3, varicella-zoster virus
C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: F27345
R:R.Davison, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: F27345
A:Molecule type: DNA
A:Residues: 1-354 <DAV>
A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CRA27950.1; PID:g60056
C:Genetics:
A:Gene: 67
C:Superfamily: herpesvirus glycoprotein D
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-354/Product: glycoprotein D #status predicted <GPD>
F:279-295/Domain: transmembrane #status predicted <TMN>
F:33, 47, 67, 116/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 8.7%; Score 162.5; DB 1; Length 354;
Best Local Similarity 22.4%; Pred. No. 1e-05;
Matches 72; Conservative 53; Mismatches 111; Indels 85; Gaps 15;
23 IIVTGNHVSARIDDDHIVTVAPRPR---ATIOLEFEMPQGRP-HKPYSGTVKVRFSDI 78

```

[illegible]



C;Genetics:  
A;Gene: 73  
C;Superfamily: pseudorabies virus glycoprotein gp63  
C;Keywords: glycoprotein; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-424/Product: glycoprotein gp63 #status predicted <MAT>  
F;320-336/Domain: transmembrane #status predicted <TM1>  
F;35,67,78,121,131,236,307/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.2%; Score 153.5; DB 1; Length 424;  
Best Local Similarity 23.8%; Pred. No. 7.1e-05;  
Matches 70; Conservative 41; Mismatches 126; Indels 57; Gaps 12;

QY 1 MASLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRENTIQL-QLFFMPG 59  
DB 1 MAKLTGMSAALLSNAICSTAIYRGEHSMYLNASSEFAVYPTQOSLVVGHLLFLDG 60  
QY 60 QR-PHKYSYGTVRVAPRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSSAN 116  
DB 61 QRLPTNYSGLIEL-IHYNSSVCYTVIQIISVSCPRVANNAFRSLKHTSKHYDYFR 119  
QY 117 RLTPGPHFKLTIRNPNDSGMFYIVRLDDTKPEIDVFAIQLSVYQFANTAATRLGLS 176  
DB 120 VNASVEINLLNITKQPTDSGAYILRVKLDHA-PTADVFGVSFAFYVDL----- 167  
QY 177 KASCRTPGLPTVQ-----LEAYLRTEESWRNMQAVATEATTS---AEATTPPTVATSA 229  
DB 168 KSKTVPDPMPETQTVEPTTSYVSTYDVTDTETESTSTQOAMTSTQTP-SATWG 226  
QY 230 SELEAEHFTFWLENGVDHVEPPANENSVTVRLG-----TMSPTL 271  
DB 227 TQLTTE-----LPTNE-----TVVIGQALLCHWFQFTRVPTL 260

RESULT 5  
VGBB63  
glycoprotein gp63 - suid herpesvirus 1  
C;Species: suid herpesvirus 1  
A;Note: host Sus scrofa domestica (domestic pig)  
C;Date: 31-Mar-1998 #sequence\_revision 31-Mar-1998 #text\_change 16-Jul-1999  
C;Accession: A29012  
R;Petrovskis, E.A.; Timmins, J.G.; Post, L.E.  
J. Virol. 60, 185-193, 1986  
A;Title: Use of lambda-gt11 to isolate genes for two pseudorabies virus glycoproteins w/  
A;Reference number: A93021; MUID:96308235; PMID:3018284  
A;Accession: A29012  
A;Molecule type: DNA  
A;Residues: 1-350 <P>  
A;Cross-references: GB:M14336; NID:G334055; PIDN:AAC35204.1; PID:G334056  
A;Experimental source: strain Rice  
C;Superfamily: pseudorabies virus glycoprotein gp63  
C;Keywords: glycoprotein; transmembrane protein  
F;11-27/Domain: transmembrane #status predicted <TM1>  
F;286-308/Domain: transmembrane #status predicted <TM2>  
F;56,73,153,256,262,275/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 143.5; DB 1; Length 350;  
Best Local Similarity 22.8%; Pred. No. 0.00037;  
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;

QY 8 LALLAATLAPGAM--GVITGNHVSARIDDDHIVIVAPRENTIQLQLFFMPGQPHKP 65  
DB 15 LLLAALTALTPRVGVLPFGAGVSVHAGSALVLPDAPNLITDGLLFLGSPSPSN- 73  
QY 66 YSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSSAN-----NRL 118  
DB 74 YSGRVEL-LRLDPRKACVTRYEAAYDLCPRVHHEAFRGCLRKEEPLARRASAAVERRL 132  
QY 119 TGPHPFKLTIRNPNDSGMFYIVRLDDTKPEIDVFAIQLSVYQFANTAATRLGL----- 174  
DB 133 -----LFVSRPAPPDAGSVLRVRVNGT---TDLFVLITLV-----PPRGRPHHP 174  
QY 175 --YSKASCRTPGLPTVQLEAYLRTEESWRNMQAVV--ATEATTS-----AEATTPTV 224

C;Genetics:  
A;Gene: 73  
C;Superfamily: pseudorabies virus glycoprotein gp63  
C;Keywords: glycoprotein; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-424/Product: glycoprotein gp63 #status predicted <MAT>  
F;320-336/Domain: transmembrane #status predicted <TM1>  
F;35,67,78,121,131,236,307/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.2%; Score 153.5; DB 1; Length 424;  
Best Local Similarity 23.8%; Pred. No. 7.1e-05;  
Matches 70; Conservative 41; Mismatches 126; Indels 57; Gaps 12;

QY 1 MASLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRENTIQL-QLFFMPG 59  
DB 1 MAKLTGMSAALLSNAICSTAIYRGEHSMYLNASSEFAVYPTQOSLVVGHLLFLDG 60  
QY 60 QR-PHKYSYGTVRVAPRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSSAN 116  
DB 61 QRLPTNYSGLIEL-IHYNSSVCYTVIQIISVSCPRVANNAFRSLKHTSKHYDYFR 119  
QY 117 RLTPGPHFKLTIRNPNDSGMFYIVRLDDTKPEIDVFAIQLSVYQFANTAATRLGLS 176  
DB 120 VNASVEINLLNITKQPTDSGAYILRVKLDHA-PTADVFGVSFAFYVDL----- 167  
QY 177 KASCRTPGLPTVQ-----LEAYLRTEESWRNMQAVATEATTS---AEATTPPTVATSA 229  
DB 168 KSKTVPDPMPETQTVEPTTSYVSTYDVTDTETESTSTQOAMTSTQTP-SATWG 226  
QY 230 SELEAEHFTFWLENGVDHVEPPANENSVTVRLG-----TMSPTL 271  
DB 227 TQLTTE-----LPTNE-----TVVIGQALLCHWFQFTRVPTL 260

RESULT 5  
VGBB63  
glycoprotein gp63 - suid herpesvirus 1  
C;Species: suid herpesvirus 1  
A;Note: host Sus scrofa domestica (domestic pig)  
C;Date: 31-Mar-1998 #sequence\_revision 31-Mar-1998 #text\_change 16-Jul-1999  
C;Accession: A29012  
R;Petrovskis, E.A.; Timmins, J.G.; Post, L.E.  
J. Virol. 60, 185-193, 1986  
A;Title: Use of lambda-gt11 to isolate genes for two pseudorabies virus glycoproteins w/  
A;Reference number: A93021; MUID:96308235; PMID:3018284  
A;Accession: A29012  
A;Molecule type: DNA  
A;Residues: 1-350 <P>  
A;Cross-references: GB:M14336; NID:G334055; PIDN:AAC35204.1; PID:G334056  
A;Experimental source: strain Rice  
C;Superfamily: pseudorabies virus glycoprotein gp63  
C;Keywords: glycoprotein; transmembrane protein  
F;11-27/Domain: transmembrane #status predicted <TM1>  
F;286-308/Domain: transmembrane #status predicted <TM2>  
F;56,73,153,256,262,275/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 143.5; DB 1; Length 350;  
Best Local Similarity 22.8%; Pred. No. 0.00037;  
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;

QY 8 LALLAATLAPGAM--GVITGNHVSARIDDDHIVIVAPRENTIQLQLFFMPGQPHKP 65  
DB 15 LLLAALTALTPRVGVLPFGAGVSVHAGSALVLPDAPNLITDGLLFLGSPSPSN- 73  
QY 66 YSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSSAN-----NRL 118  
DB 74 YSGRVEL-LRLDPRKACVTRYEAAYDLCPRVHHEAFRGCLRKEEPLARRASAAVERRL 132  
QY 119 TGPHPFKLTIRNPNDSGMFYIVRLDDTKPEIDVFAIQLSVYQFANTAATRLGL----- 174  
DB 133 -----LFVSRPAPPDAGSVLRVRVNGT---TDLFVLITLV-----PPRGRPHHP 174  
QY 175 --YSKASCRTPGLPTVQLEAYLRTEESWRNMQAVV--ATEATTS-----AEATTPTV 224

DB 175 TPSSADECR-----FWV-----GSHWDSLRVVDPAEDAVFTTTPPIEPETTPAPP 221  
QY 225 TATSAS-----ELEAEHFTFWLENGVDHVEPPANENSVTVRLGTMSPTLIGTVVA 277  
DB 222 RGTGATPEPRSEDEEDS-----EGATTANTVPFGTILDANGTM---VLNASVVSRL 271  
QY 278 AVVSATIG-----LVIVISIVTRNNCTPHRKLDTVSQDDDEERSQTRRESRK 323  
DB 272 AANAATAGAGPGKIAMVLGPTIVVLLIFLGVCACAPRCARGIASTGRDPGAARSTFR 331

RESULT 6  
S35785  
glycoprotein I - bovine herpesvirus 1  
C;Species: bovine herpesvirus 1  
C;Date: 09-Jun-1994 #sequence\_revision 12-May-1995 #text\_change 24-Nov-1999  
C;Accession: S35785  
R;Audonnet, J.  
submitted to the EMBL Data Library, June 1993  
A;Reference number: S35782  
A;Accession: S35785  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-380 <AUD>  
A;Cross-references: EMBL:Z23068; NID:G312185; PIDN:CAA80605.1; PID:G312189  
C;Superfamily: pseudorabies virus glycoprotein gp63  
C;Keywords: glycoprotein

Query Match 7.5%; Score 140.5; DB 2; Length 380;  
Best Local Similarity 25.3%; Pred. No. 0.00072;  
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;

QY 1 MASLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRENTIQL-QLFFMP 58  
DB 1 MRCLLLMVVLAAARAP--ARSLVYRGEAVGLRADGPVAFVHP-TDATLALRGSLIFLE 57  
QY 59 GQRP-HKPYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVKVTEYTSASNR 117  
DB 58 HOLPAGRRYNGTVEL-LRYHAAGDCFVMLQTTAFASCPRVANNAFRSLHADTRPARSER 116  
QY 118 LTGP--PHFPKLTIRNPNDSGMFYIVRL-----DDTKPEIDVFAIQLSVYQFA----- 166  
DB 117 RASAAVENHVLFSIARPPIDPSGLYFLRVGYGTAGSERRRDRVPLAFAVHSFGPDGP 176  
QY 167 -----NTAATRGLYSKASCRTFGL-----PTVQLEAYLRTEESWR 201  
DB 177 EAAATPAPSQRSPASGLTSSASLYDRALARSQAPPPRAPPRAARAGRRPERVDE 236  
QY 202 NMQAYVATEATTTSAEATTPPTVATSA 230  
DB 237 TTEVEAATFAGSAPALTTPPAGPTASPA 265

RESULT 7  
JQ2352  
glycoprotein I - turkey herpesvirus  
N;Alternate names: ORF 7 protein  
C;Species: turkey herpesvirus  
C;Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 07-May-1999  
C;Accession: JQ2352  
R;Zelnik, V.; Dartell, R.; Audonnet, J.C.; Smith, G.D.; Riviere, M.; Pastorek, J.; Ross  
J. Gen. Virol. 74, 2151-2162, 1993  
A;Title: The complete sequence and gene organization of the short unique region of herp  
A;Reference number: JQ2346; MUID:94014999; PMID:8409940  
A;Accession: JQ2352  
A;Molecule type: DNA  
A;Residues: 1-356 <ZEL>  
C;Keywords: glycoprotein; transmembrane protein  
F;5-27/Domain: transmembrane #status predicted <TM1>  
F;277-293/Domain: transmembrane #status predicted <TM2>  
F;154,167,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7% Score 125; DB 2; Length 356;  
Best Local Similarity 22.2%; Pred. No. 0.012;  
Matches 85; Conservative 54; Mismatches 174; Indels 70; Gaps 17;

QY 6 GTALLAATLAPF---GAMGIVITGNHVSARIDDD-HIVIVAPRPEATIQLPFMPQ 60  
DB 5 GIACWCVTLILFCIIKTCQVVRGSLVSTTVQSAQIAFFGIDTVNVLGKVLFMGQ 64  
QY 61 RPKPVSIGTVRAFRSDITNQCQELSEERFENCITRRSSVFGCKVTEYTSASNLGT 120  
DB 65 YLEIISGTMEL-LKNQANRCYSIAHATYADCPITSSVFRGCRDAVYVTRPHSRI-- 121  
QY 121 PPHF-----FKLIRNRPNDSGMVFYVIVELDTKPIDVFAIQLSVQYQFANTAARGLY 175  
DB 122 --HPQVNGLLTIEBPMDSGLIYRTSIDFNKS-----DYATISFVCMG 168  
QY 176 SKASCRT-----FGLPTVQ---LEAYLRTEESWRNW---QAYVATEATTSAEAT 220  
DB 169 SSGSCSNPRQKVSDEMCIPIHVNIAPEVYLTLVHGLPYGDLTLQIRKDMITT---APT 225  
QY 221 PTPVATASASELEAHEHTFPLWENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVVAUV 280  
DB 226 YRTIRRTVNEGLTAKTSPDIDLNATL-PLPI---SNVT---DYNMVIWRRVALRIY 278  
QY 281 SATIGLVIVISVTRNMCTPHRKLDTVSQDDEERSQTRRESKFG--PMVACEINKGADQ 338  
DB 279 AYLVIATIALIIVT--VCSAHR-----GSCSRRRIYVIGNEPTTLTSITNGNQ 326  
QY 339 DSELVELVAIVNSALSPPDSIK 361  
DB 327 EKETXNV-----PSDISDAELLE 344

## RESULT 8

T03146  
probable glycoprotein A8 - alphalaphine herpesvirus 1  
C:Species: alphalaphine herpesvirus 1  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 08-Oct-1999  
C:Accession: T03146  
R:Essner, A.; Pflanz, R.; Fleckenstein, B.  
J. Virol. 71, 6517-6525, 1997  
A>Title: Primary structure of the alphalaphine herpesvirus 1 genome.  
A:Reference number: Z14840; MUID:97404659; PMID:9261371  
A:Accession: T03146  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-683 <ENS>  
A:Cross-references: EMBL:AF005370; NID:g2337967; PIDN:AAC58098.1; PID:g2338014

Query Match 6.7% Score 124.5; DB 2; Length 683;  
Best Local Similarity 21.9%; Pred. No. 0.032;  
Matches 72; Conservative 46; Mismatches 108; Indels 103; Gaps 14;

QY 11 LAATLAPGAMGIVITGNHVSARIDDDHIVIVAPRPE--ATIQQLQFMPGQRPKPYSG 68  
DB 341 ITATISPRGEIKNVITGNF--SRNANLSLAPLSSKGKEYAGVLQA-FEPSTRPP---PG 394  
QY 69 TVRVAFRSDITNQCQELSEERFENCITRRSSVFGCKVTEYTSASNLGTGPHFPKLT 128  
DB 395 TVAPGILSTTAN-----FETSTNKSPTVT-----PTPAKLTSTPPGLTNTL 435  
QY 129 IRNRPNDSGMVFYVIVELDTKPIDVFAIQLSVQYQFANTAARGLYSKASC----- 180  
DB 436 LLTAGEHNSG-----IGSTLEPLTVSVQVQLTPSSPRDTSITLVIKLITVPDQKTV 488  
QY 181 -----RVFGLPTVOLEAYLRTEESWRNWQAYVATEATTSAEATTPPVTATASEL 232  
DB 489 SPSLVTCRTSTLIVSWTHFSRGGSPK-----PQTAAKTSEASLPPLITTT----- 538  
QY 233 EAEHFTFPLWENGVDHYEPTPAN-----ENSNVTV-----RLGTMSPTLIGV 274  
DB 539 -----PIPTNTEKQSIFASSTVSDVTTFTGDDVNTVGTWSPS---I 577

QY 275 TAAVVSATIGLVIVISVTRNMCTPHRK 303  
DB 578 TQTLPITPFTSGRQYIVV----GCCTLNAR 602

RESULT 9  
RWMS74  
T-cell surface glycoprotein CD4 precursor - mouse  
N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A02110; A36038; A39893; A39955; I54564; I69018; A47642  
R:Tourville, B.; Gorman, S.D.; Field, E.H.; Hunkapiller, T.; Parnes, J.R.  
Science 234, 610-614, 1986  
A>Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells  
A:Reference number: A02110; MUID:87018845; PMID:3094146  
A:Accession: A02110  
A:Molecule type: mRNA  
A:Residues: 1-457 <TOU>  
A:Cross-references: GB:M13816; NID:g192070; PIDN:AAA37267.1; PID:g309112  
R:Rittman, D.R.; Gethner, S.N.  
Nature 325, 453-455, 1987  
A>Title: Unusual intron in the immunoglobulin domain of the newly isolated murine CD4  
A:Reference number: A26038; MUID:87115821; PMID:3027575  
A:Accession: A26038  
A:Molecule type: mRNA  
A:Residues: 1-457 <LIT>  
A:Cross-references: GB:X04836; NID:g50353; PIDN:CAA28539.1; PID:g50354  
R:Gorman, S.D.; Tourville, B.; Parnes, J.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987  
A>Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.  
A:Reference number: A39893; MUID:88041159; PMID:2823269  
A:Accession: A39893  
A:Molecule type: DNA  
A:Residues: 1-25, 'E', 27-457 <GOR>  
A:Cross-references: GB:M17080; GB:J03003; NID:g192515; PIDN:AAA37402.1; PID:g387124  
R:Maddon, P.J.; Moliniaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.;  
Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987  
A>Title: Structure and expression of the human and mouse T4 genes.  
A:Reference number: A39955; MUID:88097446; PMID:3501122  
A:Accession: A39955  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 25-457 <MAD>  
A>Note: the cited GenBank accession number, J03564, is not in release 101.0  
R:Parnes, J.R.; Hunkapiller, T.  
Immunol. Rev. 100, 108-127, 1987  
A>Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the im  
A:Reference number: I54564; MUID:88152875; PMID:3326818  
A:Accession: I54564  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-457 <RES>  
A:Cross-references: GB:M36850; NID:g198670; PIDN:AAA39401.1; PID:g198671  
A:Accession: I69018  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 208-318 <RE2>  
A:Cross-references: GB:M36851; NID:g198672; PIDN:AAA39402.1; PID:g554183  
R:Classon, B.J.; Teagaratso, J.; Kirszbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.;  
Immunogenetics 23, 129-132, 1986  
A>Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.  
A:Reference number: A47642; MUID:86166694; PMID:3082751  
A:Accession: A47642  
A:Molecule type: protein  
A:Residues: 27-43 <CLA>  
C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells  
C:Genetics: 18/1, 74/1, 128/1, 207/1, 319/1, 386/1, 425/3, 448/2  
A:Introns: 18/1, 74/1, 128/1, 207/1, 319/1, 386/1, 425/3, 448/2  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane pr  
F1-26/Domain: signal sequence #status predicted <SIG>  
F:27-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>

```

Db      27 TVSLVSNFVDGALG-----PDGVNEDLLILG-----ELRFVGQVPHTTY   69
QY      66 YSGTVRAFRSDITNQCYQLSEERFENCSTRSSSVFGCKVTXYTFSASNRLTGPPHPF  125
DB      70 YDGGVEL-NHYPNGHKCPRVHVVTTCAPRRPAFAFALCRATDSITHS-----PAYPT  121
QY     126 -----KLTRNPRPNDGMFYIVIRLDDTKBEIDVFALQISVVQFANTATRGLYS  176
DB     122 LELNLAAQQPLLRYORATRDYAGVYLVLRVMVGDAPN-ASLFVLGMAIAAEGTLAYNGSAY-  179
QY     177 KASCRTFGCLPTVQLEAYLAITEESWRWNOAVATEATTSAEATT-PPPVTAATSASELEAE  235
DB     180 -GSCDPKLLIFS--SAPRLAPASYQPAPNQASTPSTTTTPSTIPAPTIAP--QAS  234
QY     236 HFTFPMLE-----NGVDHYEPTPEANNSNVTVLGTMSPTLIGVTVAAVYSATGLVIVI  290
DB     235 TTFPFPGDKFPQPGVGNHEPS-----NAFPAETDRSYALTVTQIIQAI PASIIALVFLG  290
QY     291 SIYTRNMCTPHRKLDITVSQDEERSQTRRESRKFQPMVACEINKGA  336
DB     291 SCI----CFIHR-----CQRYSRSRRPISPQMPTGISCAVNEAA  327

RESULT 11
A61162
glycoprotein I - equine herpesvirus 1 (strain Abl) (fragment)
C:Species: equine herpesvirus 1
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 23-Aug-1996
C:Accession: A61162
R;Elton, D.M.; Bonas, W.A.; Killington, R.A.; Meredith, D.M.; Halliburton, I.W.
Am J. Vet. Res. 52 1252-1257, 1991
A>Title: Location of open reading frames coding for equine herpesvirus type-1 glycoprot
A;Reference number: A61162; MUID:92027094; PMID:1656822
A;Accession: A61162
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A;Residues: 1-149 <BLU>
A>Note: translation of the nucleotide sequence is not complete
C;Superfamily: pseudorabies virus glycoprotein gp63
C;Keywords: glycoprotein
P;27,59,70,113,123/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          6.0%; Score 111.5; DB 2; Length 149;
Best Local Similarity 23.3%; Pred.No.0.052;
Matches 35; Conservative 33; Mismatches 69; Indels 13; Gaps 5;

QY      2 ASLGLTALLAANTLAFPGANGIVITGNHSARIDDDHHIVIVAPEATIQL-QLFEMPGQ  60
DB      2 AAILLSMAICST-----AITYRGHSMYLNASSEFAVYPTDSLVLVGHILLFDGQ  53
QY     61 R-PHKPYSGTVRAFRSDITNQCYQLSEERFENCSTRSSSVFGC--KVTEYTFASNR  117
DB     54 RLFTPNYSGLTIE-IHNYSVVCTVIQTTSYSGCPVANNAFRSLHTSKVHDIFRV  112
QY    118 LTGPFPFKLTIRNPNDGMFYIVIRLD  147
DB    113 NASVEINVLNITKPQDTSGAILRVKLD  142

RESULT 12
C39725
hypothetical protein (WS11 5' region) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Feb-1997
C:Accession: C39725
R;Samuels, M.E.; Schedl, P.; Cline, T.W.
Mol. Cell. Biol. 11, 3584-3602, 1991
A>Title: The complex set of late transcripts from the Drosophila sex determination gene
A;Reference number: A39725; MUID:91260708; PMID:1710769
A;Accession: C39725
A;Status: preliminary
A:Molecule type: mRNA
A;Residues: 1-355 <SAM>
```

C:Genetics:  
A:Gene: SGD:TUP1; ABR2; SFL2  
A:Cross-references: SGB:S0000680; MIPS:YCR084C  
A:Map position: 3R  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: transcription regulation  
F:97-118/Region: glutamine-rich  
F:181-198/Region: glutamine-rich  
F:409-422/Region: threonine-rich  
F:439-472/Domain: WD repeat homology <WD1>  
F:522-556/Domain: WD repeat homology <WD3>  
F:572-605/Domain: WD repeat homology <WD4>

Query Match 5.8%; Score 108; DB 2; Length 713;  
Best Local Similarity 27.3%; Pred. No. 0.76; Indels 46; Gaps 11;  
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPVSGTVRVAFRSDITNQCYQLSERFENCNTHRSSSVFGCKVTEYTFSS---AS 115  
DB 86 QRDHQIASLTVOORQQQQQQQQVQHLQQQQ-QQLAASASVFPVAQPPATTSATATPAA 144  
QY 116 NRUTGPPHPFKLIRNRPNDSPNMFYIVRLDDTKPEI-----D 154  
DB 145 NNTTGSFSAF--PVOASRPNLVGS-----QLPTTLFPVSSNAQQQLPQQQLQQQLQQQ 197  
QY 155 VFALQISVYOFANTAATRGILYKASCCTFGLPVQL-EAYLRTESWRNQAQV-----A 208  
DB 198 QPPQVQSVAPLSNTAING---SPTSKEKTTLPVKAPESLTKETEPENNNTSKINDTGA 254  
QY 209 TEATTTTSABAT--TPTPVATSAASEAEHFTFPWLENGVDHYEPTP 253  
DB 255 TTATTTTATEIKPKEEDATPAS-LHQHYLVFYNQR-ANHSKPIP 299

RESULT 14  
Q0B388  
glycoprotein I precursor - human herpesvirus 2 (strain 333)  
C:Species: human herpesvirus 2  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
C:Accession: A05246  
R:Hodgman, T.C.; Minson, A.C.  
Virology 153, 1-11, 1986  
A:Title: The herpes simplex virus type 2 equivalent of the herpes simplex virus type 1  
A:Reference number: A05246; MUID:86291145; PMID:3016980  
A:Accession: A05246  
A:Molecule type: DNA  
A:Residues: 1-372 <D0D>  
A:Cross-references: EMBL:M14886; NID:g330311; PIDN:AAA45861.1; PID:g330313  
C:Genetics:  
A:Map position: 0.919-0.927  
C:Superfamily: herpesvirus US7 protein  
C:Keywords: glycoprotein; transmembrane protein  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-372/Domain: glycoprotein I #status predicted <GPI>  
F:255-282/Domain: transmembrane #status predicted <TMN>  
F:156,169,175,243/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.7%; Score 107; DB 1; Length 372;  
Best Local Similarity 21.0%; Pred. No. 0.4;  
Matches 74; Conservative 51; Mismatches 163; Indels 64; Gaps 13;

QY 3 SLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPAITQL-----QLFF 56  
DB 5 SLQG-LAILGLWVC---ATGLVVRGPTVSL-VSDSLVDAGAVGQGFVEEDLRVFGELHF 59  
QY 57 MPGQRPHKPYSGTVRVAFRSDITNQCYQLSERFENCNTHRSSSVFGCKVTEYTFSSN 116  
DB 60 VQAQVPHNTYDGIIEILFHYPLGNHCPRVHVHVTLTACPRPVAFTLCRSTHAHS--- 116  
QY 117 LRTGPPHPF-----KLTIRNPRNDSGMFYIVRLDD-TKEPIDVFAIQISVYQFA 166  
DB 117 ----PAYPTLELGLARQPLLRVTRATRDYAGLYVLRVWVGATNANSLFVLGVALS----A 168

QY 167 NTAATRGYSKASCTFGLPTVQLEAYLRTESWRWQAYVATEATTSAEATTPPTVA 226  
Db 169 NGTFVYNGSDYGSCDPAQLP-----FSAPRLGPFSSVYTPGASRTPPRT 212  
QY 227 TSASELEAEHFTFPWLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGL 286  
Db 213 TTPSSSRPPTPAG-----DTGTPAPASGERAPPNSTRSASESRHRLTVAQVIQIAIPA 267  
QY 287 VIVISIVTRN-MCTPHRKLDTVSQDDERSQTRRESKFGP-MVACEINKGA 336  
Db 268 SIIAFVFLGSCICFIHRC-----QRRYRRPRGQIYNPGGVSCAVNEAA 310

RESULT 15  
F43674  
US7 protein - human herpesvirus 2 (strain HG52)  
C;Species: human herpesvirus 2  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 26-Aug-1999  
C;Accession: F43674  
R;McGeoch, D.J.; Moss, H.W.M.; McNab, D.; Frame, M.C.  
J. Gen. Virol. 68, 19-38, 1987  
A;title: DNA sequence and genetic content of the HindIII 1 region in the short unique co  
utionary comparisons:  
A;Reference number: A43674; MUID:87111457; PMID:3027242  
A;Accession: F43674  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-372 <MCG>  
A;Cross-references: EMBL:X04798; NID:G59900; PIDN:CAA28485.1; PID:G59906  
C;Superfamily: herpesvirus US7 protein

Query Match 5.3%; Score 99; DB 2; Length 372;  
Best Local Similarity 21.0%; Pred. No. 1.8; Indels 64; Gaps 13;  
Matches 74; Conservative 49; Mismatches 165;  
QY 3 SLGCTLALLAATAPGAMGIVITGNHVSARIDDDHIVIVAPPEATIQL-----QLPF 56  
Db 5 SLOG-LAILGLWC---ATGLVVRGTVSL-VSDSLVDAGAVGPGQFVEEDLRFVGFELHF 59  
QY 57 MPQQRPHKPYSGTVRVAFRSDITNQCQVELSEERFENCNTHRSSVFGCKVTEYTFASN 116  
Db 60 VQAQVPHNTYDGIIELFHYPLGNHCPRVVHVTLTACPRRPAVAFTLCRSTHAHS--- 116  
QY 117 RLTGPPHP-----KLTIRPNRNDGMPFVIVRLDD-TKEPIDVPAIQLSVYQFA 166  
Db 117 ----PAYTLELGLARQILLRVRTATRDYAGLYLVYVWGSATNASRFLGVALS----A 168  
QY 167 NTAATRGYSKASCTFGLPTVQLEAYLRTESWRWQAYVATEATTSAEATTPPTVA 226  
Db 169 NGTFVYNGSDYGSCDPAQLP-----FSAPRLGPFSSVYTPGASRTPPRT 212  
QY 227 TSASELEAEHFTFPWLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGL 286  
Db 213 TTPSSSRPPTPAG-----DTGTPAPASGEIAPPNSTRSASESRHRLTVAQVIQIAIPA 267  
QY 287 VIVISIVTRN-MCTPHRKLDTVSQDDERSQTRRESKFGP-MVACEINKGA 336  
Db 268 SIIAFVFLGSCICFIHRC-----QRRYRRPRGQIYNPGGVSCAVNEAA 310

Search completed: March 8, 2004, 02:45:23  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 01:22:05 ; Search time 27 seconds  
(without alignments)  
698.126 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863  
Sequence: 1 MASLIGTLLAATLAPFGA.....VELVAIVNPSALSFPDSIKM 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162.5	8.7	354	1 VGLI_VZVD	P09258 varicella-z
2	155	8.3	353	1 VGLI_SWVD	Q04547 simian vari
3	153.5	8.2	424	1 VGLI_HSVB	P18553 equine herp
4	143.5	7.7	350	1 VGLI_PVRI	P07646 pseudorabie
5	140.5	7.5	380	1 VGLI_HSVB	Q08102 bovine herp
6	115	6.2	457	1 CD4_MOUSE	P06332 mus musculu
7	113.5	6.1	390	1 VGLI_HSV11	P06487 herpes simp
8	108	5.8	713	1 TUP1_YEAST	P16649 saccharomyc
9	107	5.7	372	1 VGLI_HSV23	P06764 herpes simp
10	99	5.3	372	1 VGLI_HSV2H	P13291 herpes simp
11	99	5.3	1036	1 P200_MYCPN	P75211 mycoplasma
12	98.5	5.3	1140	1 YD76_MYCPN	P75405 mycoplasma
13	98	5.3	503	1 CD44_RAT	P26051 rattus norv
14	98	5.3	507	1 YG46_YEAST	P53301 saccharomyc
15	97	5.2	1220	1 SLN1_YEAST	P39928 saccharomyc
16	95.5	5.1	797	1 VGLI_HSVB	P28968 equine herp
17	93.5	5.0	907	1 VGP3_EBV	P03200 Epstein-bar
18	93	5.0	1051	1 TFIA_MOUSE	Q64127 mus musculu
19	91	4.9	836	1 NOT3_YEAST	P06102 saccharomyc
20	91	4.9	3178	1 Y889_CABEL	Q09624 caenorhabdi
21	90.5	4.9	547	1 ICA3_HUMAN	P32942 homo sapien
22	90.5	4.9	785	1 CPTA_CABEL	Q17758 caenorhabdi
23	90.5	4.9	1429	1 YC31_SCHPO	Q14065 schizosacch
24	88.5	4.8	344	1 ULB6_HCMVA	P16833 human cytom
25	88	4.7	376	1 FOS_FUGRU	P53450 fugu rubrip
26	86.5	4.6	1037	1 YHIV_ECOLI	P37637 escherichia
27	86	4.6	430	1 CINA_MYCTU	Q07731 mycobacteri
28	86	4.6	636	1 YNR6_YEAST	P53882 saccharomyc
29	85.5	4.6	503	1 PODX_MOUSE	Q9R0M4 mus musculu
30	85	4.6	376	1 MID2_YEAST	P36027 saccharomyc
31	85	4.6	583	1 CH12_CANAL	P40953 candida alb
32	85	4.6	812	1 FGRI_XENLA	P22182 xenopus lae
33	85	4.6	1718	1 RRPO_SHVX	Q04575 shallot vir

ALIGNMENTS

RESULT 1

VGLI_VZVD	STANDARD;	PRT;	354 AA.
AC P09258;			
DT 01-MAR-1989 (Rel. 10, Created)			
DT 01-MAR-1989 (Rel. 10, Last sequence update)			
DT 01-OCT-1996 (Rel. 34, Last annotation update)			
DE Glycoprotein I precursor (Glycoprotein IV) (GI) (GPV).			
GN 67.			
OS Varicella-zoster virus (strain Dumas) (VZV).			
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC Alphaherpesvirinae; Varicellovirus.			
OX NCBI_TaxID=10338;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=86306657; PubMed=3018124;			
RA Davison A.J., Scott J.E.;			
RT "The complete DNA sequence of varicella-zoster virus.;"			
RL J. Gen. Virol. 67:1759-1816(1986).			
RN [2]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=84131932; PubMed=6321154;			
RA Davison A.J.;			
RT "DNA sequence of the US component of the varicella-zoster virus genome.;"			
RL EMOB J. 2:2203-2209(1983).			
RN [3]			
RP MUTAGENESIS, AND PHOSPHORYLATION SITE.			
RX MEDLINE=94267879; PubMed=8207795;			
RA Yao Z., Grose C.;			
RT "Unusual phosphorylation sequence in the gpIV (GI) component of the varicella-zoster virus gpI-gpIV glycoprotein complex (VZV gE-gI complex).;"			
RL J. Virol. 68:4284-4211(1994).			
CC -!- SUBUNIT: HETERODIMER OF GPV AND GPI (GE).			
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC -!- SIMILARITY: TO HERPESVIRUSES GLYCOPROTEINS I, AND TO PRV GP63.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; X04370; CAA27950.1; -			
DR EMBL; X0208; CAA25032.1; -			
DR F01; F27345; VGBB67.			
DR InterPro; IPR002874; Herpes_gI.			
DR InterPro; IPR007110; IG-like.			
DR Pfam; PF01688; Herpes_gI; 1.			
DR Glycoprotein; Phosphorylation; Transmembrane; Signal.			
FT SIGNAL 1 17			
FT CHAIN 18 354			
FT DOMAIN 18 295			
FT GLYCOPROTEIN I, EXTRACELLULAR (POTENTIAL).			

O14763 homo sapien  
P13463 agrobacteri  
P08648 homo sapien  
P22314 homo sapien  
Q9erb4 rattus norv  
O88799 mus musculu  
P14644 rattus norv  
Q8ayc2 xenopus lae  
P53675 homo sapien  
P33922 escherichia  
P15379 mus musculu  
P47033 saccharomyc



```

RL Virology 189:304-316(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AB1.
RX MEDLINE=91276272; PubMed=1647359;
RA Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA Bonas W.A.;
RT "Sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equine
RT herpesvirus type-1 short unique region.";
RT Gene 101:203-208(1991).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=Kentucky D;
RX MEDLINE=91108393; PubMed=2177089;
RA Audonnet J.-C., Winslow J., Allen G., Paoletti E.;
RT "Equine herpesvirus type 1 unique short fragment encodes
RT glycoproteins with homology to herpes simplex virus type 1 gD, gI and
RT gE";
RL J. Gen. Virol. 71:2969-2978(1990).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, TO VZV GP1V,
CC AND TO PRV GP63.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; M86664; AA02508.1; -
DR EMBL; M36299; AAA66547.1; -
DR PIR; C36646; VGBEE9.
DR InterPro; IPR002874; Herpes_gI.
DR Fram; PF01688; Herpes_gI; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 GLYCOPROTEIN I.
FT DOMAIN 23 319 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 320 340 POTENTIAL.
FT DOMAIN 341 424 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 424 AA; 46392 MW; EE9BEF7DAA895806 CRC64;
Query Match 8.2%; Score 153.5; DB 1; Length 424;
Best Local Similarity 23.8%; Pred. No. 3.3e-05;
Matches 70; Conservative 41; Mismatches 126; Indels 57; Gaps 12;

Qy 1 MASLLGTALLANTLAPFGAMGIVITGNHVSARRDDHHIVIVAPRPATQL-QLFMPG 59
Db 1 MAKLTGTFSAILLISMAICSTAIIRGHEMSMYLNASFEAVYPTDQSLVVGHLFLDG 60
Qy 60 QR-PHKYSGTVRVAFRSDITNQCQQLSEERFENCTHRSSVFGC--KVTEYTFGSAN 116
Db 61 QRLPTTNYSLGIEU--IHNYSVVCYTIQTISYSCRPVANNAPRSLHKTSKYHDYFR 119
Qy 117 RLTPGPPHFKLTIRPNRPNMGFYVIRLDDTKRPIDVFAIQLSVYQFANTAATRGILYS 176
Db 120 VNASVETNVLNITKQPTDSGAILRVKLDHA-PTADVFGVSFAFVDL----- 167
Qy 177 KASCRTEGLPTVQ---LEAYLRTEESWRNMQAYVATEATTS---AEATPTPTVATSA 229
Db 168 KSKIVPDMPTQIVPEPTTSVVSPTVDYTDVITEESTSTSQQAMTSQTPT-SATWG 226
Qy 230 SELSAEHFTFPWLENGVDHYEPTPANENSNVTVRLG-----TMGPTL 271

```

```

Db      227  TQLTVE-----LPTNE-----TWVIGQEAALLCHWFQPSRVRPTL 260

RESULT 4
VGLI PRVRI
ID      VGLI PRVRI      STANDARD;      PRT;      350 AA.
AC      P07646;
DT      01-APR-1988 (Rel. 07, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      Glycoprotein GP63 precursor.
OS      Pseudorabies virus (strain Rice) (PRV).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Alphaherpesvirinae; Varicellovirus.
OX      NCBI_TaxID=103350;
RN      [1].
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86308235; PubMed=3018284;
RA      Petrovskis B.A., Timmins J.G., Post L.E.;
RT      "Use of lambda gt11 to isolate genes for two pseudorabies virus
RT      glycoproteins with homology to herpes simplex virus and varicella-
RT      zoster virus glycoproteins.";
RL      J. Virol. 60:185-193(1986).
CC      CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, AND TO VZV
CC      GPV.
CC      CC
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      CC
CC      CC EMBL; M14336; AAC35204.1; -.
CC      PIR; A29012; VGBE63.
CC      DR      InterPro: IPR002874; Herpes_g1.
CC      DR      Pfam; PF01688; Herpes_g1, 1.
CC      KW      Glycoprotein; Transmembrane; Signal.
CC      FT      SIGNAL      1      23
CC      FT      CHAIN      24      350
CC      FT      DOMAIN      24      285
CC      FT      TRANSMEM      286      308
CC      FT      DOMAIN      309      350
CC      FT      CARBOHYD      56      56
CC      FT      CARBOHYD      73      73
CC      FT      CARBOHYD      153      153
CC      FT      CARBOHYD      256      256
CC      FT      CARBOHYD      262      262
CC      FT      CARBOHYD      275      275
CC      SQ      SEQUENCE      350 AA; 36773 MW; 66AF2229EC21BEDA CRC64;
CC      CC
CC      CC Query Match
CC      Best Local Similarity      7.7%; Score 143.5; DB 1; Length 350;
CC      Matches      82; Conservative      42; Mismatches      149; Indels      87; Gaps      17;
QY      8      LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQLQFFMFGQRPKHP 65
DB      15      LLALALTLAALTRPGVGLFRGAGCVHYVAGSAVIVPGDAPNLITDGLLLFLGSPSPN- 73
QY      66      YSGTVRVAFRSDITNOCYQBSLSEPFENCETHRSSSVFVGC--KVTEYTFAS-----NRL 118
DB      74      YSRGVEL-LRLDPKRCYTRYEAAEYDLCPRVHFAFRGCLRKREPLARRASAARELL 132
QY      119      TGPHPPKLIRNPRNDSGNFVIVRLDDTKEPIDVFAIQLSVYQFANTATRGL---- 174
DB      133      -----LFSVPAPPDAGSYVLRVRNGT---TDLFVLITLV-----PPRGPHHP 174
QY      175      --YSKASCRFGLPTVLEAYLFTESWRNWAQVY--ATEATWTS-----AEATPTPV 224
DB      175      TPSSADSCR----PV-----GSHWDSLRVDPDAEDAVFTTPPPIPEPPTTAPP 221

```



```
QY 225 TATSAS-----ELEAEHFTPPMLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVTA 277
Db 222 RGTGATPERSDEEED-----EGATTATMPVPGTLDANGTM---VLNASVSVSVLL 271
QY 278 AVVSATIG-----LVVISIVTRNMCTPHRKLDTVSQDDERSQTRRESRX 323
Db 272 AAANATAGARGPKIAMVLGETIVVLLIFLGVCACARRCARGIAGTGRDGPAGARRSTR 331
RESULT 5
VGLI_HSVBS STANDARD; PRT; 380 AA.
AC 08102;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein 1.
GN Gi.
OS Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=45407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167875; PubMed=8122370;
RA Leung-Tack P., Audonnet J.F., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short
RT unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL Virology 199;409-421(1994).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS 1, TO VZV GP1V,
CC AND TO PRV GP63.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 223668; CAA80605.1; -.
DR PIR; S35785; S35785.
DR InterPro; IPR002874; Herpes_g1.
DR Pfam; PF01688; Herpes_g1; 1.
KW Glycoprotein.
FT CARBOHYD. 67 67 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 380 AA; 39910 MW; ABBELFB9B430D2BD CRC64;
Query Match
Best Local Similarity 7.5%; Score 140.5; DB 1; Length 380;
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;
QY 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATQL--QLFFMP 58
Db 1 MRCILLMMVLAARAP--ARSLYRGEAVGLRADGPAFAVHP-TDATALRGLRFLFE 57
QY 59 GQRP-HKYSCTGVVARSDITNOCYQELSEEPENCETHSVFFGCKVTEFPASNR 117
Db 58 HOLPAGRYNGTVEL-LRYHAAGDCFVMLQTTAFASPRVANNAPFSLHADTRPARSER 116
QY 118 LTGP--PHPFKLTIRNRPDNDGFMFYIVRL-----DDTKPIDVFAQLSVYQFA---- 166
Db 117 RASAAVENHVLFSATRPIDSGLYFLRVGIYGTAGSERRRDVPFLAFAVHSFGPGDP 176
QY 167 -----NTAATRGLYSKASCTFGI-----PTVQLEAVLRTEESWR 201
Db 177 EAAARTPAPRSQSRPAASGLTSSASLYDALARSPOAPPRPAPPAARAGRRPRPVRDE 236
QY 202 NQAVVATEATTSEATTPFTVATAS 230
Db 237 TTEVEAATRASAFALTTPPAGPTASPA 265
```

## RESULT 6

```
CD4_MOUSE
ID CD4_MOUSE STANDARD; PRT; 457 AA.
AC P06332;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3) (T-cell differentiation antigen L3T4).
GN CD4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87018845; PubMed=3094146;
RA Tourville B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;
RT "Isolation and sequence of L3T4 complementary DNA clones: expression
RT in T cells and brain.";
RL Science 234;610-614(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115821; PubMed=3027575;
RA Littman D.R., Gettner S.N.;
RT "Unusual intron in the immunoglobulin domain of the newly isolated
RT murine CD4 (L3T4) gene.";
RL Nature 325;453-455(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88152875; PubMed=3326818;
RA Parnes J.R., Hunkapiller T.;
RT "L3T4 and the immunoglobulin gene superfamily: new relationships
RT between the immune system and the nervous system.";
RL Immunol. Rev. 100;109-127(1987).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=88041159; PubMed=2823269;
RA Gorman S.D., Tourville B., Parnes J.R.;
RT "Structure of the mouse gene encoding CD4 and an unusual transcript
RT in brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 84;7644-7648(1987).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
RA Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
RA Gibbs R.A.;
RT "Comparative sequence analysis of a gene-rich cluster at human
RT chromosome 12p13 and its syntenic region in mouse chromosome 6.";
RL Genome Res. 8;29-40(1998).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smolus D.E.,
```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7]  
 RP SEQUENCE OF 27-43.  
 RX MEDLINE=86166694; PubMed=3082751;  
 RA Classon B.J., Tsgaratos J., Kirsbaum L., Maddox J., McKay C.R.,  
 RA Brandon M., McKenzie I.P.C., Walker I.D.;  
 RT "The L374 antigen in mouse and the sheep equivalent are  
 immunoglobulin-like.";  
 RL Immunogenetics 23:129-132(1986).  
 RN [8]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=86233454; PubMed=3086886;  
 RA Classon B.J., Tsgaratos J., McKenzie I.F.C., Walker I.D.;  
 RT "Partial primary structure of the T4 antigens of mouse and sheep:  
 assignment of intrachain disulfide bonds.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).  
 CC -!- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 receptor interaction. May regulate T-cell activation.  
 CC -!- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P06332-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Brain-specific;  
 CC IsoId=P06332-2; Sequence=VSP\_002489;  
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M36850; AAA39401.1; -  
 DR EMBL; M13816; AAA37267.1; -  
 DR EMBL; X04836; CAA28539.1; -  
 DR EMBL; M36851; AAA39402.1; -  
 DR EMBL; M17080; AAA37403.1; -  
 DR EMBL; M17078; AAA37403.1; JOINED.  
 DR EMBL; M17079; AAA37403.1; JOINED.  
 DR EMBL; AC002397; AAC36010.1; -  
 DR EMBL; BC039137; AAC39137.1; -  
 DR PIR; A02110; RWMST4.  
 DR HSSP; P01730; 1WBR.  
 DR MGD; MGI:88335; Cd4.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . . ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ISS.  
 DR InterPro; IPR000973; CD4\_TcRg.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV\_1.  
 DR PROSITE; PS0833; Ig\_Like; 1.  
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate;  
 KW Alternative splicing.  
 FT SIGNAL 1 26  
 FT CHAIN 27 457 T-CELL SURFACE GLYCOPROTEIN CD4.  
 FT DOMAIN 27 394 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 395 417  
 FT DOMAIN 418 457  
 FT DOMAIN 27 128  
 FT DOMAIN 129 207  
 FT DOMAIN 208 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 187 187  
 FT CARBOHYD 298 298  
 FT CARBOHYD 323 323  
 FT CARBOHYD 392 392  
 FT DISULFID 42 112  
 FT DISULFID 159 188  
 FT DISULFID 328 370  
 FT LIPID 418 418  
 FT LIPID 421 421  
 FT VARSPLIC 1 240  
 FT SEQUENCE 457 AA; 51296 MW; 1B1DA7527CB00F33 CRC64;  
 SQ  
 Query Match 6.2%; Score 115; DB 1; Length 457;  
 Best Local Similarity 20.6%; Pred. No. 0.055;  
 Matches 76; Conservative 54; Mismatches 133; Indels 106; Gaps 16;  
 QY 19 GAGCIVITGNHVSARIDDDHIVIVAPRPETATIQQLF---FMGPQPHKPYSTVTVAFR 75  
 DB 93 GSFPFLII--NKLKMEQSYICELENRKE-EVELWVFKVTFSPGTSLLQCSLTLLDSN 149  
 QY 76 SDIYNQCYQLSEERFENCETHRSSVYVCKKTEYTFSSANRLTGPHFPFKLTIRPNR 135  
 DB 150 SKVSN-----PLTE-----CHKKGVVSGSKV-----LSMSNLRVQ 181  
 QY 136 DSGMFFVIVRLDDTKPEIDVFAIQLSYVQFANTAATRGLYSKASCRTFGLPVQLEAYLR 195  
 DB 182 DSDPFWNCTVLDQKK---NWFGMTLSVLGFQSTAIT-AVKSGBESAEFSPLNFAB--- 233  
 QY 196 TEESWRN--WQAVYATEATTSABEATTPVATSAEELAEHPTFPWLENGVDHVEPTP 253  
 DB 234 -ENGWGLMWKA-----EKDSFFQFWISFSIKKEVSV 265  
 QY 254 ANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCTPHRKLDPVSQDDDE 313  
 DB 266 QKSTKDLKQLKETLPLTKIPQVSLQFAGSGML-----TLTLDKGLHVEVNLVWVKVAQ 321  
 QY 314 RSQTRRESRKFGPMVACEIN-----KGADQD---SELVELVAIVNPSA----- 353  
 DB 322 LNNT-----LTCEVMGPTSPKMLTLKQENAEARVSEBQKYVQVVAETGLWQCL 371  
 QY 354 LSSPDSIKM 362  
 DB 372 LSEGDVKVM 380  
 RESULT 7  
 VGLI\_HSV11  
 ID VGLI\_HSV11 STANDARD; PRT; 390 AA.  
 AC P06487;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Glycoprotein I.  
 GN GI OR US7.  
 OS Herpes simplex virus (type 1 / strain 17).  
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 CC NCBI\_TaxID=10299;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=85160822; PubMed=2984429;  
 RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;  
 RT "Sequence determination and genetic content of the short unique  
 region in the genome of herpes simplex virus type 1.";  
 RL J. Mol. Biol. 181:1-13(1985).  
 CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND



FT REPEAT 670 706 WD 7.  
FT DOMAIN 97 118 POLY-GLN.  
FT DOMAIN 181 198 POLY-GLN.  
FT DOMAIN 399 409 THR-RICH.  
FT CONFLICT 75 75 E -> A (IN REF. 1).  
FT CONFLICT 100 100 R -> Q (IN REF. 1).  
FT CONFLICT 685 685 P -> S (IN REF. 1 AND 2).  
SQ SEQUENCE 713 AA; 78307 MW; 444104AAD63CB944 CRC64;  
  
Query Match 5.8%; Score 108; DB 1; Length 713;  
Best Local Similarity 27.3%; Pred. No. 0.37; Mismatches 89; Indels 46; Gaps 11;  
Matches 62; Conservative 30;  
  
QY 60 QRPKPYSGTVAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFSS-----AS 115  
DB 86 QRDHQIASITVQOORQQQQQQVQHLQOQQ-QQLAASASVPVQAQPPATTSATAPAA 144  
  
QY 116 NRLTGPPHPFKLTIRNPRNDSGMFVIVRLDDTKPEI-----D 154  
DB 145 NTTGSPSAF-PVQASRENVLGS-----QLPTTLFVSSNAQQQLPQQLOOQOQQ 197  
  
QY 155 VFAIQLSVVQFANTAAATRGLYKASCRTEGLTVQL-EAYLRTESWRNQAIV-----A 208  
DB 198 QPPQVSVAPLNTAING---SPTSKETTLPSVKAPESTLKETEPENNTSKINDTGA 254  
  
QY 209 TEATTTSAEAT--TPTPVATASAELEAEHFTFPLENGVDHYEPTP 253  
DB 255 TTATTTTTEIKPEEDATPAS-LHQDHYLVIPNQR-AHNSKPIP 299

RESULT 9  
VGLI\_HSV23  
ID VGLI\_HSV23 STANDARD; PRT; 372 AA.  
AC P06764;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Glycoprotein I.  
GN GI OR US7.  
OS Herpes simplex virus (type 2 / strain 333).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86291145; PubMed=3016980;  
RA Hodgman T.C., Minson A.C.;  
RT "The herpes simplex virus type 2 equivalent of the herpes simplex virus type 1 US7 gene and its flanking sequences."  
RL Virology 153:1-11(1986).  
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND 2: GH, GB, GC, GD, GI, AND GE.  
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, TO VZV GP1V, AND TO PRV GP63.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M14886; AAA45861.1; --  
CC EMBL; D00026; BAA00021.1; --  
CC PIR; A05246; Q0BE88.  
CC InterPro; IPR002874; Herpes\_g1.  
CC Pfam; PF01688; Herpes\_g1; 1.  
CC Glycoprotein.  
KW CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 372 AA; 39548 MW; BB3A9E2D20D8AD3D CRC64;  
  
Query Match 5.7%; Score 107; DB 1; Length 372;  
Best Local Similarity 21.0%; Pred. No. 0.19; Mismatches 51; Indels 64; Gaps 13; Matches 74; Conservative 51;  
  
QY 3 SLIGTLALATLAAPFGAMGIVITGNHVSARIODDDHIVIVAPPEATIQI-----QLFF 56  
DB 5 SLOG-LAILGLWVC---ATGLVVRGPTVSL-VSDSLVDAGVGPQGFVEDLRFVGLHFP 59  
  
QY 57 MPQORPKPYSGTVAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFSSAN 116  
DB 60 VQAQVPTNVDYDGLIELFHYPLGNHCPRVHVVTLTACPRPAVAFTLCHSTHAHS--- 116  
  
QY 117 RLGTGPPHPF-----KLTIINRPNDGSMFVIVRLDD-TKEPIDVFAIQLSVTQFA 166  
DB 117 ---PAYTLELGLARQPLLRVTRATRDYAGLVYLVVWGSATNASLFLGVALS-----A 168  
  
QY 167 NTAATRGLYKASCRTEGLTVQL-EAYLRTESWRNQAIVATEATTSAEATTPPTVA 226  
DB 169 NGTFVNGSDYGCDAQLP-----FSAPRLGPSSVYTFGASRTPPT 212  
  
QY 227 TSASELEAEHFTFPLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVAAVVSATIGL 286  
DB 213 TTSFSSPRDTPAPG-----DTGTPAPASGERAPPNSTRSASERHRLTVAQVIOIAPA 267  
  
QY 287 VIVISIVTRN-MCTPHRKLDTVSQDDDEERSQTRRESKFGP-MVACEINKGA 336  
DB 268 SIATFVLGSCICFIHRC-----QRRYRPRGQIYNPFGVSCAVNEAA 310

RESULT 10  
VGLI\_HSV2H  
ID VGLI\_HSV2H STANDARD; PRT; 372 AA.  
AC P13291;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Glycoprotein I.  
GN GI OR US7.  
OS Herpes simplex virus (type 2 / strain HG52).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10315;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87111457; PubMed=3027242;  
RA McGeoch D.J., Moss H.W.M., McNab D., Frame M.C.;  
RT "DNA sequence and genetic content of the HindIII 1 region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons."  
RL J. Gen. Virol. 68:19-38(1987).  
RN [2]  
RP SEQUENCE FROM N.A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X04798; CAA28485.1; --  
CC EMBL; Z86099; CAB06714.1; --  
CC PIR; F43674; F43674.

DR InterPro: IPR002874; Herpes\_g1.  
DR Pfam: PF01688; Herpes\_g1; 1.  
KW Glycoprotein.  
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 372 AA; 39558 MW; COD4A22CAB16B8D7 CRC64;  
  
Query Match 5.3%; Score 99; DB 1; Length 372;  
Best Local Similarity 21.0%; Pred. No. 0.88;  
Matches 74; Conservative 49; Mismatches 165; Indels 64; Gaps 13;  
  
QY 3 SILGTLIALAATLAPGANGIVITGNHVSARIDDDHIVIVAPREAIQL-----QLFF 56  
DB 5 SLQG-LAILGLWVC---ATGLVVRGPTVSL-VSDSLVDAGAVGPGFVEEDLRVFGELHF 59  
  
QY 57 MFCQRHKPYSGTVRVAFPSDITNOCYQELSEERFENCNTHRSSVFGCKTETVFFSASN 116  
DB 60 VQAQVPHNYDGIIELFHYPLGNHCPRVVHVVTITACPRRPAVAFTLCRSTHHAHS--- 116  
  
QY 117 RLUTGPHPF-----KLTIKRNPNDSGMFYVIVRLDD-TKSPIDVFAIQLSVYQFA 166  
DB 117 ----PAYPTLELGLAQPLLRVTATRDYAGLYLVLRVWVGSATNASRFLGVALS---A 168  
  
QY 167 NTAATRGYSKASCTRTGLPTVQLEAYLRTESNRNQAYVATEATTTSAEATTTPTVTA 226  
DB 169 NGTFVYNGSDYSCDPAQLP-----FSAPRLGPSVVTYPGASRTPPT 212  
  
QY 227 TSASELEAHFFPWLNGVDHYEPTPANENSNTVRLGTMSPTLIGTVAAVVSATIGL 286  
DB 213 TTPSSFRDPTAPG-----DTGTPAPASGETAPPNSTRSASESHRLTLVAQVQIAIPA 267  
  
QY 287 VIVISIVTRN-MCTPHRKLDVTSQDDEERSQTRRSRKEGP-MWACEINKGA 336  
DB 268 SIIFAVLFGSCICFIHRC-----QRRYRPRGQIYNGVSCAVNEAA 310  
  
RESULT 11  
ID P200\_MYCPN STANDARD; PRT; 1036 AA.  
AC P75211; Q50346;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein P200.  
GN P200 OR MPN367 OR MP275.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=96257197; PubMed=8675035;  
RA Proft T., Hilbert H., Plagens H., Herrmann R.;  
RT "The P200 protein of Mycoplasma pneumoniae shows common features with  
RT the cytochrome-associated proteins HMW1 and HMW3."  
RL Gene 171:79-82(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae."  
RL Nucleic Acids Res. 24:4420-4449(1996).  
RN [3]  
RP SEQUENCE OF 641-678 FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=95075318; PubMed=7984111;  
RA Proft T., Herrmann R.;  
RT "Identification and characterization of hitherto unknown Mycoplasma

RT pneumoniae proteins.";  
RL Mol. Microbiol. 13:337-348(1994).  
CC -!- FUNCTION: PROTEIN CYTOSKELETON-ASSOCIATED WHICH COULD BE AN  
CC ACCESSORY STRUCTURAL COMPONENT IN CYTADHERENCE.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-aib.ch/announce/>  
CC or send an email to [licenses@isb-sib.ch](mailto:licenses@isb-sib.ch)).  
CC  
DR EMBL; U25989; AAC9815.1; -  
DR EMBL; AE00027; AAB95923.1; -  
DR EMBL; Z32646; CAA83569.1; -  
DR PIR; S73601; S73601.  
KW Cytoadherence; Structural protein; Repeat; Complete proteome.  
FT DOMAIN 277 280 POLY-THR.  
FT DOMAIN 300 845 PRO-RICH.  
FT DOMAIN 357 360 POLY-THR.  
FT DOMAIN 401 404 POLY-ALA.  
FT DOMAIN 718 781 3 X 6 AA REPEAT OF E-P-E-P-N-F.  
FT REPEAT 718 723 1.  
FT REPEAT 738 743 2.  
FT REPEAT 776 781 3.  
FT CONFLICT 641 641 A -> P (IN REF. 3).  
SQ SEQUENCE 1036 AA; 116915 MW; DESAEBABDD95B29 CRC64;  
  
Query Match 5.3%; Score 99; DB 1; Length 1036;  
Best Local Similarity 23.9%; Pred. No. 3.3;  
Matches 57; Conservative 29; Mismatches 90; Indels 62; Gaps 10;  
  
QY 127 LTRNPNDSGMFYVIVRLDDTKSPIDVFAIQLSVYQFAN---TAATRGYSKASCTTF 183  
DB 446 WIDNGQPQAGHYVVDFTLT-APLTVAEIELQELVNEFVTTTRETTFPSTPVF 504  
  
QY 184 G---LPTVQLEAYLRTESNRNQAYVATEATTTSEATTP---TPVTATSALEAEH 236  
DB 505 EPVVIPTVSEEQLENE-----FVSTVVSATSNPNVASTPVVT----- 546  
  
QY 237 FTTFWLENGVDHYEPTPANENSNTVRLGTMSPTLIGTV-----AAVVSATIGL 286  
DB 547 -----VELTE-TPVSLPELETVQLETPAVVTEITVTEKAVEPLVAVVEAPLAV 596  
  
QY 287 VIVISIVTRNMCTPHRKLDVTSQDDEERSQTRRSRKEG--PMVACINEKADQDSL 342  
DB 597 EPIVETST-----TLAETVEEAQVEQBSTAVAVEPALETESKATSEAAEL 643  
  
RESULT 12  
ID YD76\_MYCPN STANDARD; PRT; 1140 AA.  
AC P75405;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MPN376 (A19\_orf1140).  
GN MPN376 OR MP460.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae."  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; STRONG, TO M.PNEUMONIAE  
CC MPN375.

CC -!- SIMILARITY: IN THE CENTRAL SECTION; STRONG, TO M.PNEUMONIAE  
CC MPN374.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO M.PNEUMONIAE  
CC MPN373.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; A8000045; AAB96108.1; --  
DR PIR; S73786; S73786.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 8 28 POTENTIAL.  
FT TRANSMEM 1098 1118 POTENTIAL.  
SQ SEQUENCE 1140 AA; 130383 MW; 8FA9406C57DD8886 CRC64;  
  
Query Match 5.3%; Score 98.5; DB 1; Length 1140;  
Best Local Similarity 20.5%; Pred. No. 4;  
Matches 68; Conservative 53; Mismatches 119; Indels 91; Gaps 15;  
  
QY 71 RVAERSDITNQC-YOELSERFE-NCTHSSSVFVGVCKVTEYTFGSANRLTGPPHPFKLT 128  
Db 680 REFQKDIINYLVKVGQISTEIKVAVHKVDAMLNARKSDDKPLAS---VGTANKYGLN 736  
QY 129 IR-NRPNDSGMFYIV-----RLDDTK-----EPIDVFAIQLSVVQFAN-- 167  
Db 737 LRSNPY---TGQFYVVVDVNDLGNQRANNAKSYFYIIGLKGQSSVLRFEKQ 793  
QY 168 -----TAATRGLYSKASCTFGLPTVQLBAYLRTE-----WRNQAYVATEAVT 213  
Db 794 KLYSLESIAVDSKGLVVK-----NVSKDAIIQAKQNLKYLDTTHWNA--ALKANL 842  
QY 214 TGAETATPTPVATASALEAEAEHFPFWLGVHVEYTPANENSIVTVGLTMSPTLIG 273  
Db 843 TNAELTLPT-ASADNSAKLS-----TPNAENDE-----GFUSENVSG 878  
QY 274 VTVAAVVSATIGLIVISIVTRNM-----CTPHKLDIVSDDEERSQTRRE 320  
Db 879 SILGYVYRTGKLFUKERSVFNKEDKNLKLRLTSNFTLDKKGKLVDPSPVINOIVEE 938  
QY 321 SRKFGPMACEIKNGADQDSSELVVAIVNP 351  
Db 939 ARGYNVLVSEKGDDESKNFKITLTNP 969  
  
RESULT 13  
CD44\_RAT  
ID CD44\_RAT STANDARD; PRT; 503 AA.  
AC P26051; Q99021;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)  
DE (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte  
DE homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)  
DE (LY-24).  
GN CD44.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC STRAIN=BDIX; TISSUE=Pancreas;  
RX MEDLINE=9119152; PubMed=1707342;  
RA Guenther U., Hofmann M., Rudy W., Reber S., Zoeller M.,  
RA Hausmann I., Matzku S., Wenzel A., Ponta H., Herrlich P.,  
RT "A new variant of glycoprotein CD44 confers metastatic potential to  
RT rat carcinoma cells."

Cell 65:13-24 (1991).  
[2]  
SEQUENCE FROM N.A. (ISOFORM 1).  
Stevens J.W., Midura R.J.;  
Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to  
mucosal high endothelial venule and to types I and VI collagen.  
Probably involved in matrix adhesion, lymphocyte activation and  
lymph node homing.  
-!- SUBCELLULAR LOCATION: Type I membrane protein.  
-!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Comment=Additional isoforms seem to exist;  
Name=2; Synonyms=Long, Meta-1;  
Name=1; Synonyms=Short;  
IsoId=P26051-1; Sequence=Displayed;  
IsoId=P26051-2; Sequence=VSP\_005330;  
-!- PTM: N-glycosylated (By similarity).  
-!- PTM: O-glycosylated; contains chondroitin sulfate glycans which  
can be more or less sulfated (By similarity).  
-!- PTM: Phosphorylated; activation of PKC results in the  
dephosphorylation of Ser-467 (constitutive phosphorylation site),  
and the phosphorylation of Ser-433 (By similarity).  
-!- SIMILARITY: Contains 1 link domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M61875; AAA53532.1; --  
EMBL; M61874; AAA53534.1; --  
EMBL; U52179; AAA97915.1; --  
EMBL; U46957; AAA92920.1; --  
PIR; B38745; B38745.  
HSP; P98066; ITSG.  
InterPro; IPR001231; CD44 antigen.  
InterPro; IPR000538; Link.  
Pfam; PF00193; Xlink; 1.  
PRINTS; PR00658; CD44.  
PRINTS; PR01265; LINKMODULE.  
PRODOM; PD000918; Link; 1.  
SMART; SM00445; Link; 1.  
PROSITE; PS01241; Link; 1.  
Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;  
Proteoglycan; Signal; Alternative splicing;  
Pyroglutamate carboxylic acid.  
FT SIGNAL 1 21  
FT CHAIN 22 503 CD44 ANTIGEN.  
FT DOMAIN 22 410 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 411 431 POTENTIAL.  
FT DOMAIN 432 503 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 51 123 LINK.  
FT DOMAIN 154 162 ARG/LYS-RICH (BASIC).  
FT DOMAIN 228 410 STEM.  
FT MOD\_RES 22 22 PYROLIDONE CARBOXYLIC ACID (PROBABLE).  
FT DISULFID 56 122 BY SIMILARITY.  
FT DISULFID 80 100 BY SIMILARITY.  
FT MOD\_RES 433 433 PHOSPHORYLATION (BY PKC) (PARTIAL) (BY  
SIMILARITY).  
FT MOD\_RES 467 467 PHOSPHORYLATION (PARTIAL) (BY  
SIMILARITY).  
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 224 385 IATPPWSAHTKQKQERTQWNPISHNPEVLLQTTTRMTDID  
RNTSAHGNWQEQPOPPFNHNEYQDEBETPHATITWADP  
FT FT NSTTEAATQKSKWENWQGNPPTPSDSHVTRGTASA  
FT FT HNNPSQRTTOSQEDVSWTDFDPISHPMGQGHQTESK  
FT FT -> SDGSSMDPRGGFDIVHGSCLA (in isoform  
1).  
FT FT /FTid=VSP 005330.  
FT FT R -> S (IN REF. 2).  
FT FT  
SQ SEQUENCE 503 AA; 55945 MW; 74 74 (IN REF. 2);  
Query Match 5.3%; Score 98; DB 1; Length 503;  
Best Local Similarity 22.2%; Pred. No. 1.6;  
Matches 47; Conservative 30; Mismatches 79; Indels 56; Gaps 10;  
QY 178 ASCTRTGLPVOLEAVLRTEESW-RNWQ-----AYVATEATTSAEATTPVT 225  
DB 297 ATSTTADPNSTTTEAATQKQKFNENWQGNPPTPSDSHV-TEGTTASAHNHPQRM 355  
QY 226 ATGASELEABHFTFWLENGVDHYEP-----TPANENSVTVRLGTWSP 269  
DB 356 TTQSQE-----DVSW-----TDFDPISHPMGQGHQTESKHSNGQDSGVTTTSGPARR 405  
QY 270 TLIG-----VTVAVVSATIGLVIVISIVTRNMCTPHRKLPTVSDQDEERSQTRRGRKFGP 326  
DB 406 PQIPEWLIILASLLALLILAVCIANVSRRCQKKL--VINGNGTVEDRKPS----- 458  
QY 327 MVACEINKGADQSELVELVAIVNPSALSPD 358  
DB 459 ----ELNGEASKQEMVHL---VNKEPTETPD 483

RESULT 14  
YQ46 YEAST STANDARD; PRT; 507 AA.  
ID YQ46 YEAST  
AC P53301;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 52.8 kDa protein in BUB1-HIP1 intergenic region.  
GN YGR189C OR G7553.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=97279231; PubMed=9133739;  
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,  
RA Nombela C.;  
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm  
of Saccharomyces cerevisiae chromosome VII.";  
RL Yeast 13:357-363(1997).  
CC -!- SIMILARITY: SOME, TO YEAST UTR2.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 272974; CAA97215.1; -;  
DR EMBL; X99074; CAA67525.1; -;  
DR PIR; S64507; S64507.  
DR HSSP; P23904; LAJK.  
DR GernOnline; 141501; -;  
DR SGD; S0003421; CRH1.  
DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.  
DR GO; GO:0000131; C:incipient bud site; IDA.  
DR InterPro; IPR008985; ConA like lec.gl.  
DR InterPro; IPR000757; Glyco\_hydro\_16.

DR Pfam; PF00722; Glyco\_hydro\_16; 1.  
KW Hypothetical protein.  
FT DOMAIN 63 66 POLY-SER.  
FT DOMAIN 301 310 POLY-SER.  
FT DOMAIN 345 357 POLY-SER.  
FT DOMAIN 387 391 POLY-SER.  
FT DOMAIN 467 470 POLY-SER.  
SQ SEQUENCE 507 AA; 52757 MW; 7D7B61F57ABA942C CRC64;  
Query Match 5.3%; Score 98; DB 1; Length 507;  
Best Local Similarity 21.7%; Pred. No. 1.6;  
Matches 54; Conservative 46; Mismatches 95; Indels 54; Gaps 9;  
QY 136 DSGNFYVIVRLDDTKEPIDVFAIQLSVYQFANTAAATGLYSKASCTRTGLPTVOLEAYLR 195  
DB 278 DGGSIY-GRYDQAGDAVLANGGSISSSTSSSTVSSSASTVSSSVSSVSSSASST 335  
QY 196 TEESWRNQAVATEATTSAEATTPVTATGASELEABHFTFWLENGVDHYEPTPAN 255  
DB 336 VSSS---VSSVSSSVSSSVSSSTSPSSSTATSKTLASSVT---TSSSISSPE-KQSS 388  
QY 256 ENSNVTVRLGTWSPTLIGVT-VAAVVSATIGLVIVISIVTRNMCTPHRKLDTVSQD--- 310  
DB 389 SSSKKTVASSTSSSIISSTKTPATVST-----TRSTVAPTQSSVSSDSFVQ 438  
QY 311 -----DEERSQTRRGRKFGPMV-----ACEINKGADQD-----S 340  
DB 439 DKGAVATSSNDVTSTTQISSKYTSTIQSSSSEASTNSVQISNGADLAQSLPREGKLF 498  
QY 341 ELVELVAIV 349  
DB 499 VLVALALL 507

RESULT 15  
SLN1 YEAST STANDARD; PRT; 1220 AA.  
ID SLN1 YEAST  
AC P39528;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Osmolarity two-component system protein SLN1 (EC 2.7.3.-).  
GN SLN1 OR YPD2 OR YII47C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / YPH1;  
RX MEDLINE=94024010; PubMed=8211183;  
RA Ota I.M., Varshavsky A.;  
RT "A yeast protein similar to bacterial two-component regulators.";  
RL Science 262:566-569(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=97313266; PubMed=9169870;  
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,  
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,  
RA Harris D.E., Hornslett T., Hunt S., Jagels K., Jones M., Lye G.,  
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,  
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";  
RL Nature 387:84-87(1997).  
RN [3]  
RP MUTAGENESIS OF HIS-576 AND ASP-1144.  
RX MEDLINE=94239498; PubMed=8183345;  
RA Maeda T., Wurgle-Murphy S.M., Saito H.;  
RT "A two-component system that regulates an osmosensing MAP kinase  
RT cascade in yeast.";  
RL Nature 369:242-245(1994).  
CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:30:36 ; Search time 125 Seconds  
(Without alignments)  
913.741 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863

Sequence: 1 MASLLGTALLAATLAPFGA.....VELVAIVNPSALSSPDSIKM 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	362	Q67645	Q67645 gallid herp
2	164	8.8	420	Q39309	Q39309 equine herp
3	162.5	8.7	354	Q98VNI	Q98VNI human herpe
4	156	8.5	420	Q35525	Q35525 equine herp
5	156	8.4	384	Q29357	Q29357 feline herp
6	154	8.3	384	Q26931	Q26931 feline herp
7	152	8.2	355	Q266L5	Q266L5 turkey herp
8	151	8.1	355	Q84LU8	Q84LU8 turkey herp
9	149	8.0	355	Q69287	Q69287 turkey herp
10	148	7.9	364	Q291P1	Q291P1 canine herp
11	148	7.9	364	Q41525	Q41525 canine herp
12	147	7.9	259	Q91335	Q91335 canine herp
13	145	7.8	370	Q86789	Q86789 feline herp
14	144	7.7	355	Q84LU7	Q84LU7 turkey herp
15	140.5	7.5	369	Q91E39	Q91E39 phocid herp
16	139.5	7.5	382	Q39505	Q39505 bovine herp

17	139	7.5	355	12	Q8JLU9	Q8JLU9 turkey herp
18	130.5	7.0	366	12	Q99F67	Q99F67 suid herpes
19	125	6.7	356	12	Q88524	Q88524 turkey herp
20	124.5	6.7	383	12	Q36401	Q36401 alcelaphine
21	123	6.6	355	12	Q67638	Q67638 gallid herp
22	115	6.2	433	11	Q50554	Q50554 mus musculu
23	115	6.2	457	11	Q61396	Q61396 mus musculu
24	112	6.0	387	4	Q86XK7	Q86XK7 homo sapien
25	112	6.0	582	5	Q81R74	Q81R74 drosophila
26	108.5	5.8	319	5	Q9W3S7	Q9W3S7 drosophila
27	108.5	5.8	475	16	Q8FQU1	Q8FQU1 corynebacte
28	108	5.8	381	5	Q81R75	Q81R75 drosophila
29	106.5	5.7	235	12	Q9DY19	Q9DY19 bovine herp
30	103	5.5	982	5	Q9VIG1	Q9VIG1 drosophila
31	102.5	5.5	378	4	Q96H15	Q96H15 homo sapien
32	102.5	5.5	4498	5	Q9W2Z3	Q9W2Z3 drosophila
33	100.5	5.4	2114	5	Q81Q18	Q81Q18 drosophila
34	99	5.3	279	5	Q9U474	Q9U474 caenorhabdi
35	99	5.3	1011	5	Q9NHZ9	Q9NHZ9 helicoverpa
36	98	5.3	264	11	Q80XS5	Q80XS5 mus musculu
37	98	5.3	862	11	Q8K0T1	Q8K0T1 mus musculu
38	98	5.3	1262	11	Q80T73	Q80T73 mus musculu
39	97.5	5.2	1711	5	Q9W1X4	Q9W1X4 drosophila
40	96.5	5.2	814	13	Q91897	Q91897 xenopus lae
41	96	5.2	877	4	Q9H3Q6	Q9H3Q6 homo sapien
42	96	5.2	878	4	Q9H3Q7	Q9H3Q7 homo sapien
43	96	5.2	957	4	Q14651	Q14651 homo sapien
44	96	5.2	1217	4	Q9UKW9	Q9UKW9 homo sapien
45	95.5	5.1	332	12	Q8V0L9	Q8V0L9 equine herp

#### ALIGNMENTS

#### RESULT 1

Q67645 ID Q67645 PRELIMINARY; PRT; 362 AA.  
AC Q67645;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glycoprotein 1.  
OS Gallid herpesvirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.  
OX NCBI\_TaxID=10386;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA challenge strain;  
RX MEDLINE=97033380; PubMed=8879127;  
RA Wild M.A., Cook S., Cochran M.;  
RT "A genomic map of infectious laryngotracheitis virus and the sequence and organization of genes present in the unique short and flanking regions."  
RL Virus Genes 12:107-116(1996).  
DR EMBL; U28832; AAC55101.1; .  
DR InterPro; IPR002874; Herpes\_g1.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF01688; Herpes\_g1; 1.  
SQ SEQUENCE 362 AA; 39750 MW; F530C1AAF7CC6BA5 CRC64;

Query Match 100.0%; Score 1863; DB 12; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.9e-162;  
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPRTATLQQLFFMPGQ	60
Db	1	MASLLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPRTATLQQLFFMPGQ	60
QY	61	RPHKPYGTVRVAFRSDITNCQVQELSEERFENCTHRSSSVFVGCCKTETFFSASNLITG	120
Db	61	RPHKPYGTVRVAFRSDITNCQVQELSEERFENCTHRSSSVFVGCCKTETFFSASNLITG	120

```
QY 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
Db 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
QY 181 RTFGLPTVQLEAYLRTESWRNQAAYVATEATTTSAEATTPPTVATTSASELEAEHFTFP 240
Db 181 RTFGLPTVQLEAYLRTESWRNQAAYVATEATTTSAEATTPPTVATTSASELEAEHFTFP 240
QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGVTVAAVVSATIGLVIVISVITNMCTP 300
Db 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGVTVAAVVSATIGLVIVISVITNMCTP 300
QY 301 HRKLDTVSODDERSOTRESRKFGPMVACEINKGADQDSELVELVAIVNPSALSPDSI 360
Db 301 HRKLDTVSODDERSOTRESRKFGPMVACEINKGADQDSELVELVAIVNPSALSPDSI 360
QY 361 KM 362
Db 361 KM 362

RESULT 2
O39309 PRELIMINARY; PRT; 420 AA.
AC O39309;
DT 01-JAN-1998 (TrEMBLrel. 05, Last created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Counterpart of HSV-1 gene US7 and VZV gene 67.
GN 73.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RX MEDLINE=98264497; PubMed=9603335;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-4.";
RL J. Gen. Virol. 79:1197-1203 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030027; AAC59593.1; -.
DR F01; T42616; T42616.
DR InterPro; IPR002874; Herpes_g1.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF01688; Herpes_g1; 1.
SQ SEQUENCE 420 AA; 45710 MW; 1EB944825DF3D852 CRC64;

Query Match 8.8%; Score 164; DB 12; Length 420;
Best Local Similarity 25.5%; Pred. No. 2.1e-06;
Matches 96; Conservative 50; Mismatches 143; Indels 88; Gaps 21;

QY 10 LLAATLAPFGAMGIVTGNHVSARIDDDHIVIVAPREAIQI-QLFPMGQR-PHKPYS 67
Db 13 LLAISMCM--ATAIYRGEHMSVYNLNASSEFAVYPKDKSLVYVGHMLFDGRLPTNYS 69
QY 68 GTVAVAFRSIDITNQCQELSEERFENCTHRSSSVFVGC--KVTEYTFASNRLTGPFP 125
Db 70 GLIEL-IHNSYRGCVSIQTIYESCFRVANAFRSLHKTSHNNQDVFHNTSVETNV 128
QY 126 KLTIRNPRNDGMYFVIVRLDDTKPIDVFAIQLSVYQF-ANTAATRGLYSKASCTFG 184
Db 129 LLNITRPOPADSGAIILURVKLNHA-PTADVFGSAFYDLOQSNVPEPPTAKEPSNVT 187
QY 185 LPTVQLEAYLRTESWRNQAAYVATEATTTSAEATTPPTVATTSASELEAEHFTFP 241
Db 188 RTPAPANTSK-----TGSNTTSSQSTWLYTTP-----RPA 221
```

```
QY 242 LENGVDHYEPTPANEN--SNVTVRL-----GTMSP-----LIGTVAAVVSATIG----- 285
Db 222 LET--HLTTPANETVWSGDTAMLCHGRSPSTAVPTIYMLLGLT-----GNLPE 269
QY 286 --LVIVISVITRMCTPHKLDTVS-----ODDEERSQIRRESF-KFGPMV-----ACEIN 333
Db 270 DVLLIEDSEILR---TPPPKQPTTSRTGDDPKQNTSTSPKSRNKIVAMVVIPTACVLM 326
QY 334 KGADQDSELVELVAIVN 350
Db 327 -----LLLVVVGAIIN 337

RESULT 3
Q98VN1 PRELIMINARY; PRT; 354 AA.
ID Q98VN1
AC Q98VN1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORE67.
OS Human herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10335;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-32;
RX MEDLINE=21109116; PubMed=11162813;
RA Faga B., Maury M., Bruckner D.A., Grose C.;
RT "Identification and Mapping of Single Nucleotide Polymorphisms in the
RT Varicella-Zoster Virus Genome.";
RL Virology 280:11-6 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-32;
RA Cole N.L., Faga B.P., Grose C.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-VIA;
RA Grose C., Faga B.;
RT "Identification and mapping of single nucleotide polymorphisms in the
RT varicella-zoster virus genome.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF314221; AAK19249.1; -.
DR EMBL; AF325441; AAK01055.1; -.
DR InterPro; IPR002874; Herpes_g1.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF01688; Herpes_g1; 1.
SQ SEQUENCE 354 AA; 39373 MW; 3F01739F3AFC6B08 CRC64;

Query Match 8.7%; Score 162.5; DB 12; Length 354;
Best Local Similarity 22.4%; Pred. No. 2.2e-06;
Matches 72; Conservative 53; Mismatches 111; Indels 85; Gaps 15;

QY 23 IVITGNHVSARIDDDHIVIVAPRE---ATIQQLFFMPGQR-PHKPYSGTVRVAFRSDI 78
Db 21 LIFKGDHVSQVNSLSLILPMQNDNVTIKGLVIGEQLPTGNTVSGTLELLY-ADT 79
QY 79 TNQCQELSEERFENCTHRSSSVFVGCQVT---BYTFSASNRLTGPFPFKLTIRNPRN 135
Db 80 VAFCFRSVQVIRYDGCPRITSAFISCRKSHWHYGNSTDRISTEPDAGVMKTKPGIN 139
QY 136 DSGMFPYIVRLDDTKPIDVFAIQLSVY-----QFANTAATRGLYSKAS 179
Db 140 DAGVYVLLVRLDHSR-TDGFILGVNVYTAGSHNIHGVITSPSLQNGYSTRALFQA- 197
QY 180 CRTGLEPT-----VQLEAYLRTES-WRNQAAYVATEATTTSAEATTPPT 225
Db 198 -RLCDLPATPKSGTSLFOHMLDLRAGKSLDNLFWLH-EDVVTETKSVKVG- 248
QY 226 ATSASELEAEHFTFPWLENGVDHYEPTPAN-----ENS-----NVTVRLGTSPTLIGTV 276
```

```
Db 249 -----IEN-----HVYPTDSTLPEKSLNDPENLLI-----LIPIVA 281
QY 277 AAVVSATGLVIVISVTRNM 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 SVMILTAMVIVISVGRRI 302

RESULT 4
OS5525 PRELIMINARY; PRT; 420 AA.
AC O55525;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH20;
RX MEDLINE=98185635; PubMed=9524947;
RA Damiani A.M., Matsumura T., Yokoyama N., Maeda K., Miyazawa T.,
RA Kai C., Mikami T.;
RT "Nucleotide sequences of glycoprotein I and E genes of equine
RT herpesvirus type 4.";
RL J. Vet. Med. Sci. 60:219-225 (1998).
DR EMBL; AB005749; BAA25021.1; -.
DR InterPro; IPR002874; Herpes GI.
DR Pfan; PF01688; Herpes GI; 1.
SQ SEQUENCE 420 AA; 45696 MW; 1F64F5E179C2D916 CRC64;

Query Match 8.5%; Score 158; DB 12; Length 420;
Best Local Similarity 25.5%; Pred. No. 7.3e-06;
Matches 96; Conservative 49; Mismatches 144; Indels 88; Gaps 21;

QY 10 LLAATLAPFAGMIVITGNHVSARIDDDHIVIVAPRPEATLQLQLFMPGQR-PHKPYS 67
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 LLAISNC---ATAIYRGHSMYLNASSEFAVYPKDKSLVVGVHMLFDGRLFTNYS 69
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 GTVAVFRSDITNQCQLSEERFENCTHRSSSVFVGC--KVTEYTFASNRLGPPHPF 125
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 GLIEL-IHNSYRGCSYQIQTISYSCPRVANNAPRSC LHKTSNHNQDYFHTVSTVETNV 128
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 126 KLITRNPRNDGMPFYVIVRLDDTKEPIDVFAIQLSVYQF-ANTAATRGLYSKASCRFTG 184
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 LLNITWQPQADSGAILRVKLNHA-PTADVGVSAFVYDLSQNTVPPEVPTAKEPSNVPT 187
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 185 LPTVQLEAYLRTEESWRNQAVVATEATTSABAT---TPTPTVATSAASELEAEHFTFPW 241
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 RTPAPAPANTSTK-----TGSNTTSQSSTWLYTPP-----RPA 221
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 LENGVDHVEPTANEN--SNVTVRL-----GTMSPT-----LIQVTVAAVVSATIG---- 285
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 LET---HLTAPANETVSGDTAMLCHGRPRSTAVPTIYMLLGLT-----GNLPE 269
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 --LVIVISVTRNMCTPRKLDVTSQDDE-----ERSOTRRESR-KFGPMV-----ACIN 333
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 DVLIIEDSEILR---TPPKPQOTTSSRTGDAFKQTNSTSPSKRNKIVAMVVIPTACVIM 326
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 334 KGADQDSELVELVAIVN 350
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 -----LLLVVVGALIN 337

RESULT 5
Q69357 PRELIMINARY; PRT; 384 AA.
ID Q69357
AC Q69357;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
```

```
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Homologue of HSV-1 GI.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2620;
RX MEDLINE=95266277; PubMed=7747442;
RA Willemse M.J., Strijdeven I.G., van Schooneveld S.H.,
RA van den Berg M.C., Sondermeijer P.J.;
RT "Transcriptional analysis of the short segment of the feline
RT herpesvirus type 1 genome and insertional mutagenesis of a unique
RT reading frame.";
RL Virology 208:704-711 (1995).
DR EMBL; D42113; BAA07692.1; -.
DR InterPro; IPR002874; Herpes GI.
DR Pfam; PF01688; Herpes GI; 1.
SQ SEQUENCE 384 AA; 43009 MW; 670C0893CEAAB9DC CRC64;

Query Match 8.4%; Score 156; DB 12; Length 384;
Best Local Similarity 23.3%; Pred. No. 9.9e-06;
Matches 75; Conservative 48; Mismatches 141; Indels 58; Gaps 12;

QY 5 LGTLALATLAPFAGMIVITGNHVSARIDDDHIVIVAPRPE-ATIQQLFFMPGQ-RP 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MSSIAFIYILMAIGTVYGVGRDHSVLDVTSSTGFIYPTLENFTYHGLIFLDQQLP 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 HKPYSGTVAVFRSDITNQCQLSEERFENCTHRSSSVFVGC--KVTEYTFASNRLTG 120
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 VNNYNGTLEI-IHYNHSCYKIVQIEYSSCPVRNNAFRSCLHKTSMHQYDQLSINTS 119
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 PPHPFKLITRNPRNDGMPFYVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 VETGMLLTITSPQVEDGGIYALVRFNHNK-ADVFGLSVFVYSP----DTRGRRHAD- 173
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 RTFGLPTVQLEAYLRTEESWRNQAVVATEATTSABATTPPTVATSAASELEAEHFTFP 240
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 -----ENLNGSILLTPS--SMETVYKVTPIYDHMTTQT-----TSNKSMESE----P 216
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 WLENGVDHVEPTANENSNVTVRL-----GTMSPTLIQVTVAAVVSAT- 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 SNTSISCHTFQNDPNEGELYTHLLIAGNITYDDVMVDGTTLPRLIDGLNLSVTSF 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 284 -----IGLVIVISI 292
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 KNETTQKWTDPDKVGFVIVISI 298
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
Q66931 PRELIMINARY; PRT; 384 AA.
ID Q66931
AC Q66931;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Feline herpesvirus type 1 GI gene precursor.
GN GI.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B927;
RX MEDLINE=96357074; PubMed=8764058;
RA Mijnes J.D.F., van der Horst L.M., van Anken E., Horzinek M.C.,
RA Rottier P.J.M., De Groot R.J.;
RT "Biosynthesis of glycoproteins E and I of Feline herpesvirus: gE-gI
RT interaction is required for intracellular transport.";
RL J. Virol. 70:5466-5475 (1996).
```

DR EMBL; X98448; CAA67077.1; -.  
DR InterPro: IPR002874; Herpes\_gi.  
DR Pfam: PF01688; Herpes\_gi; 1.  
KW SIGNAL.  
FT SIGNAL 1 18 POTENTIAL.  
SQ SEQUENCE 384 AA; 43019 MW; C9ACE1429E496DC1 CRC64;  
  
Query Match 8.3%; Score 154; DB 12; Length 384;  
Best Local Similarity 22.3%; Pred. No. 1.5e-05;  
Matches 73; Conservative 52; Mismatches 133; Indels 70; Gaps 13;  
  
QY 5 LGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPE-ATIQQLFFMPGQ-RP 62  
Db 1 MSIAPIVILMAIGTVGVGRGDVSLHVDTSFGVIYPTLENFTIYGHILFLDQPLP 60  
  
QY 63 HKPYGTAVRSDITNOCYQELSEERFENCTHRSSSVFVGC--KVTEYFASNLRTG 120  
Db 61 VANNYNGTLEI-THYHSHSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHGYDQLSINTS 119  
  
QY 121 PHPPFKLTIRNPRPNDGPFYIVIRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASC 180  
Db 120 VETGMLLTITSPKMDGGIYALRVFNNHKK-ADVGLSVFYISF-----DTRGHRHDAE 174  
  
QY 181 RTFG--LPTVQ--LEAYLRTEESRWQAY---VATEATTSAEATTPVATTSASELEA 234  
Db 175 NLNGEILITPSPMETTVKV-----NTPYDHNVTTQTSNKSWESEPSNTSISC----- 223  
  
QY 235 EHFTFPWLENGVDHYEFTPANENSNVTVRL-----GTMSPTLIGTVVAA 278  
Db 224 -----HTQNDPNEGETLYTHLLNAGNITVDDMVDGTTLKPLIDMGLNL 270  
  
QY 279 VVSAT-----IGLVIVISI 292  
Db 271 SVTSSFKNETQKWPDRKVGIVIVISI 298

RESULT 7  
Q9E6L5 PRELIMINARY; PRT; 355 AA.  
AC Q9E6L5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Glycoprotein I.  
GN MDV095 OR GIS.  
OS Turkey herpesvirus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Marek's disease-like viruses.  
OX NCBI\_TaxID=10390;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MD5;  
RX MEDLINE=20392152; PubMed=10933706;  
RA Tulman E.R., Afonso C.L., Lu Z., Zaak L., Rock D.L., Kutish G.F.,  
RT "The genome of a very virulent Marek's disease virus."  
J. Virol. 74:7980-7988 (2000).  
[2]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=MD5;  
RA Tulman E.R., Afonso C.L., Lu Z., Zaak L., Rock D.L., Kutish G.F.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=X;  
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;  
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three  
Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations  
in the Glycoprotein L-encoding Gene in Some vv+MDVs."  
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RL  
DR EMBL; AF243438; AAM97723.1; -.  
DR EMBL; AY129990; AAM97723.1; -.  
DR InterPro: IPR002874; Herpes\_gi.  
DR Pfam: PF01688; Herpes\_gi; 1.

SQ SEQUENCE 355 AA; 40568 MW; D23A6DA126ACDB00 CRC64;  
  
Query Match 8.2%; Score 152; DB 12; Length 355;  
Best Local Similarity 22.9%; Pred. No. 2.1e-05;  
Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;  
  
QY 19 GAMGIVITGNHVSARIDDDHIVIVAPRPE-ATIQQLFFMPGQRPKPKSGTVVAFRSD 77  
Db 15 GIWSIVVTGTSVLTSDQSALVAFCGLDKMNVRGQLFLGDQTRTSSYTGTTETI-LKWD 73  
  
QY 78 ITNOCYQELSEERFENCTHRSSSVFVGCVKTEYFASNLRTGPPHPFK---LTIRNRP 134  
Db 74 EBYKCYSLVATSYMDCPAIDATVFRGCRDAAVYAQPHDRV--OPPEKGTLLRIIVEPRV 131  
  
QY 135 NDSGMFVIVRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASCRTFGLPTVQLEAYL 194  
Db 132 SDTGSYIIRVALAG-RNMSDIFRMAVIIRS-----SKSWACNHSASSFQAHKCI 179  
  
QY 195 RTEE--SWRNQAYVATEATTSAEATTPVATTSASELEASHFTFPWLENGVDHYEPT 252  
Db 180 RYVDRMAFENYLGHVGNLLDSDSELHAYINIPQGIS-TDINIITPFYDNGSTIYSPT 238  
  
QY 253 P---ANENSNVTVRLGT-MSPTLIGTVVAAV--SATIGLVIV---ISIVTRNMCTPHRK 303  
Db 239 VFNLFNNSHVDAMNSTGMWNTVLKYLPLRIYFSTMVILCIILALAIYLCERCSPHRR 298  
  
QY 304 LDTVSQDDERSQTRRESKFGPMVACEINKGADQDSOLVE 344  
Db 299 I----YIGEPSRDE-----APLIITSANVESFOYDYNKKE 328

RESULT 8  
Q8JUL8 PRELIMINARY; PRT; 355 AA.  
AC Q8JUL8;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Glycoprotein I.  
GN GIS.  
OS Turkey herpesvirus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Marek's disease-like viruses.  
OX NCBI\_TaxID=10390;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N;  
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;  
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three  
Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations  
in the Glycoprotein L-encoding Gene in Some vv+MDVs."  
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RL  
DR EMBL; AY129992; AAM97725.1; -.  
DR InterPro: IPR002874; Herpes\_gi.  
DR Pfam: PF01688; Herpes\_gi; 1.  
SQ SEQUENCE 355 AA; 40584 MW; 58DB07B04193C0A0 CRC64;

Query Match 8.1%; Score 151; DB 12; Length 355;  
Best Local Similarity 23.2%; Pred. No. 2.6e-05;  
Matches 79; Conservative 53; Mismatches 167; Indels 42; Gaps 14;  
  
QY 19 GAMGIVITGNHVSARIDDDHIVIVAPRPEATIQI-QLFFMPGQRPKPKSGTVVAFRSD 77  
Db 15 GIWSIVVTGTSVLTSDQSALVAFCGLDKMNVRGQLFLGDQTRTSSYTGTTETI-LKWD 73  
  
QY 78 ITNOCYQELSEERFENCTHRSSSVFVGCVKTEYFASNLRTGPPHPFK---LTIRNRP 134  
Db 74 EBYKCYSLVATSYMDCPAIDATVFRGCRDAAVYAQPHDRV--OPPEKGTLLRIIVEPRV 131  
  
QY 135 NDSGMFVIVRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASCRTFGLPTVQLEAYL 194  
Db 132 SDTGSYIIRVALAG-RNMSDIFRMAVIIRS-----SKSWACNHSASSFQAHKCI 179

```
QY 195 RTEE--SWRNQOYVATEATTTSABATTPVATASASELBAEHFTFPMLENGVDHYEPT 252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 RVDRMAFENYLIGHVGNLSDSELHAIYITPQIS--TDINIITTFYDNGSCTIYSPT 238
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 P---ANENSVTVRLGT--MSPTLIGVTVAAV--SATIGLVIV---ISIVTRNCTPHRK 303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 VFNLFNNSHVDAMNSTMNTVLKYTLPLRVIYFSTMVLCIIALAIYLCVRCRSPHR 298
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 LDTVSQDDEERSQTRRESKFGPMVACINKGADQDSSELVE 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 I-----YIGEPRDE-----APLITSAVNESFOYDYNKE 328
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
Q69287 PRELIMINARY; PRT; 355 AA.
AC Q69287;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glycoprotein I.
GN US7 OR GIS.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95133166; PubMed=7831788;
RA Brnovakis P., Velicer L.F.;
RT "The Marek's disease virus (MDV) unique short region:
RT alphaherpesvirus-homologous, fowlpox virus-homologous, and MDV-
RT specific genes.";
RL Virology 206:324-338 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RB1B;
RA Parcellis M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; L22174; AAA64968.1; -.
DR EMBL; AY129994; RAM9727.1; -.
DR InterPro; IPR002874; Herpes_GI.
DR Pfam; PF01688; Herpes_GI; 1.
DR Signal.
KW SIGNAL.
FT CHAIN
SQ SEQUENCE 355 AA; 40579 MW; 40CD25034E4EAE6F CRC64;

Query Match 8.0%; Score 149; DB 12; Length 355;
Best Local Similarity 23.3%; Pred. No. 3.9e-05;
Matches 80; Conservative 54; Mismatches 163; Indels 46; Gaps 15;

QY 19 GAWGVITGNHVSARIDDDHIVIVAPR---PEATIQQLFFMPGQPHKPSGTGVRAFR 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 GWSIVYTTSTV--LSTDQSLAVAFRGLDKAVNVRGQLLFGDQTRTSSVTGTETI-LK 71
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 SDITNQCYOELSEERFENCNTHRSSSVFGCKVTEYTFSSASRLTGPPHPFK---LTIRNP 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 WDEEYKCYSLVHATSMDCPAIDATVFRGCRDAVVAQPHGRV--QPFPEKGTLLRIVEP 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 RNDSGMFVIVRLDDTKPEIDVFAIQLSVYGFANTATRGLYSKASCTRTGLFTVQLEA 192
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 RVSDTGYSYIRVSLAG--RNMSDIFRMVWIIRS-----SKSWACNHSASSFQAHK 177
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 YLRTEE--SWRNQOYVATEATTTSABATTPVATASASELBAEHFTFPMLENGVDHYE 250
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 CIRYVDRMAFENYLIGHVGNLSDSELHAIYITPQIS--TDINIITTFYDNGSCTIYS 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 251 PTP---ANENSVTVRLGT--MSPTLIGVTVAAV--SATIGLVIV---ISIVTRNCTPH 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 237 PTVFNLFNNSHVDAMNSTMNTVLKYTLPLRVIYFSTMVLCIIALAIYLCVRCRSPH 296
QY 302 RKLDTVSQDDEERSQTRRESKFGPMVACINKGADQDSSELVE 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 RRI-----YIGEPRDE-----APLITSAVNESFOYDYNKE 328
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q9YFAL PRELIMINARY; PRT; 364 AA.
AC Q9YFAL;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to HSV-1 gI.
OS Canine herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=37110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CSL-2;
RX MEDLINE=98185505; PubMed=9524817;
RA Tyack S.G., Studdert M.J., Johnson M.A.;
RT "Nucleotide sequence of canine herpesvirus homologues of herpes
RT simplex virus type 1 US2, US3, glycoproteins I and E, US8.5 and US9
RT genes.";
RL DNA Seq. 7:365-368 (1997).
DR EMBL; U49380; AAC67213.1; -.
DR GO; GO:0005529; P:sugar binding; IEA.
DR InterPro; IPR002874; Herpes_GI.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01688; Herpes_GI; 1.
DR SMART; SM00034; CLECT; 1.
DR Signal.
SQ SEQUENCE 364 AA; 41999 MW; B9BC33BE7FB913A9 CRC64;

Query Match 7.9%; Score 148; DB 12; Length 364;
Best Local Similarity 27.3%; Pred. No. 5e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTALLAATLAPGAGMIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQR-P 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16 LLITMFLPILFLFLYGVNGFYKGTYSIMFLNTSSGFSFPDDKFIUSGRLLDQHL 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 HKPYSGTVRVAFRSDITNQCYOELSEERFENCNTHRSSSVFVGC--KVTEYTFSSASRLTG 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 VNNYSGTIEFI---HFNNSCYTVYQIEYFSCPRIFNNAFRSCLKXVKHESQLRINS 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 PPHPFKLITIRNPNDSGMPYIVRLDDTKPEIDVFAIQLSVYQF 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 IENGVLLEITNPKPNDSGVYFIRVQLENNK--TDVFGIPAFIYSF 175
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
O41525 PRELIMINARY; PRT; 364 AA.
AC O41525;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CUS7 (Glycoprotein I).
GN GI.
OS Canine herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=37110;
RN [1]
RP SEQUENCE FROM N.A.
RA Haanes E.J., Tomlinson C.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=YPI11MU;
RX MEDLINE=98455388; PubMed=9784067;
RA Nishikawa Y., Xuan X., Otsuka H.;
RT "Identification and characterization of the glycoprotein E and I genes
RT of canine herpesvirus.";
RN Virus Res. 56:77-92(1998).
RP [3]
RC STRAIN=Australian;
RA Reubel G.H., Pekin J., Webb-Wagg K., Hardy C.M.;
RT "Nucleotide sequence of glycoprotein genes B, D, G, H and I, thymidine
RT kinase and protein kinase genes and gene homologue UL 24 of an
RT Australian isolate of canine herpesvirus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84223; AAB67059.1;
DR EMBL; AF003729; BAA33765.1;
DR EMBL; AF361076; AAK51063.1;
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR002874; Herpes_gi.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01688; Herpes_gi; 1.
DR SMART; SM00034; CLECT; 1.
SQ SEQUENCE 364 AA; 42068 MW; 08BC32EF7FB913A8 CRC64;

Query Match 7.9%; Score 148; DB 12; Length 364;
Best Local Similarity 27.3%; Pred. No. 5e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQQR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYISMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKVSKHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKPNDGSGVFIRVQLENKK--TDVFGIPAFIYSF 175
SQ SEQUENCE 370 AA; 41568 MW; 2A816ECB37456A77 CRC64;

Query Match 7.8%; Score 145; DB 12; Length 370;
Best Local Similarity 22.0%; Pred. No. 9.6e-05;
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;

QY 5 LGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRE-ATTQQLFFMPQQR-P 62
Db 1 MSSIAFIYILMAIGTVYGVYRGDVSILHVDTSSTSGFVYPTLENFTIYGHILFDQDPLP 60
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 61 VNNYNGTLEI-IHYNHSSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHOYDQLSINTS 119
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQFANTAARGLYSKASC 180
Db 120 VETGMLLTITSPKMGDGIYALRVFRFNHKK-ADVFGLSVFVYSF---DTRGHRHHADE 174
QY 181 RTFG--LFTVQ--LEAYLFTESWRNQQAY---VATEATTTSAEATTPFPVATASALEA 234
Db 175 NINGELLTPSPMETYKVV-----NTPYIDEMVTTQTTSNKSMSESEPTNTSISC----- 223
QY 235 EHFTFPWLENGVDHYEPTPANENSVTVRL-----GTMSFTLIGTVAA 278
Db 224 -----HTFQNDPNEGETLYTHLINIAGNITYDDMVMMDGTTLKPLLI----- 264
QY 279 VVSATIGLVIVISIVTRNMCTPHRKLDT 306
Db 265 ----DMGLNLSVTSPFKN--GNHAKMDT 286

RESULT 14
QYUL07
ID Q8JLU7
AC Q8JLU7;

Query Match 7.9%; Score 147; DB 12; Length 259;
Best Local Similarity 27.3%; Pred. No. 3.8e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQQR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYISMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKVSKHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKPNDGSGVFIRVQLENKK--TDVFGIPAFIYSF 175
SQ SEQUENCE 259 AA; 29965 MW; 66C7A9A957FF7E0F CRC64;

Query Match 7.9%; Score 147; DB 12; Length 259;
Best Local Similarity 27.3%; Pred. No. 3.8e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQQR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYISMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKVSKHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKPNDGSGVFIRVQLENKK--TDVFGIPAFIYSF 175
SQ SEQUENCE 259 AA; 29965 MW; 66C7A9A957FF7E0F CRC64;
```

```
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQQR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYISMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKVSKHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKPNDGSGVFIRVQLENKK--TDVFGIPAFIYSF 175
SQ SEQUENCE 370 AA; 41568 MW; 2A816ECB37456A77 CRC64;

Query Match 7.8%; Score 145; DB 12; Length 370;
Best Local Similarity 22.0%; Pred. No. 9.6e-05;
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;

QY 5 LGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRE-ATTQQLFFMPQQR-P 62
Db 1 MSSIAFIYILMAIGTVYGVYRGDVSILHVDTSSTSGFVYPTLENFTIYGHILFDQDPLP 60
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 61 VNNYNGTLEI-IHYNHSSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHOYDQLSINTS 119
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQFANTAARGLYSKASC 180
Db 120 VETGMLLTITSPKMGDGIYALRVFRFNHKK-ADVFGLSVFVYSF---DTRGHRHHADE 174
QY 181 RTFG--LFTVQ--LEAYLFTESWRNQQAY---VATEATTTSAEATTPFPVATASALEA 234
Db 175 NINGELLTPSPMETYKVV-----NTPYIDEMVTTQTTSNKSMSESEPTNTSISC----- 223
QY 235 EHFTFPWLENGVDHYEPTPANENSVTVRL-----GTMSFTLIGTVAA 278
Db 224 -----HTFQNDPNEGETLYTHLINIAGNITYDDMVMMDGTTLKPLLI----- 264
QY 279 VVSATIGLVIVISIVTRNMCTPHRKLDT 306
Db 265 ----DMGLNLSVTSPFKN--GNHAKMDT 286

RESULT 14
QYUL07
ID Q8JLU7
AC Q8JLU7;

Query Match 7.9%; Score 147; DB 12; Length 259;
Best Local Similarity 27.3%; Pred. No. 3.8e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQQR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYISMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYSQELSEERFENCNTHRSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKVSKHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFVIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKPNDGSGVFIRVQLENKK--TDVFGIPAFIYSF 175
SQ SEQUENCE 259 AA; 29965 MW; 66C7A9A957FF7E0F CRC64;
```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein 1.
GN GlS.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK;
RA Parcells M.S., Shamblin C.E., Dienglelcz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vMDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129993; AAM97726.1; -
DR InterPro: IPR002874; Herpes_gi.
DR Pfam: PF01688; Herpes_gi; 1.
DR SEQUENCE 355 AA; 40599 MW; 316993F97143C1E0 CRC64;

Query Match      7.7%; Score 144; DB 12; Length 355;
Best Local Similarity 22.6%; Pred. No. 0.00011;
Matches 77; Conservative 55; Mismatches 167; Indels 42; Gaps 14;

QY 19 GAMGIVITGNHVSARIDDDHIVIVAPRP-ATIQQLFFMPQGRPHKPYSGTVRVARFSD 77
DB 15 GIWSRYVTGTSVLTSDQALVAFCLDKWVNRQLFLGQDTTSYTGTEI-LKWD 73

QY 78 ITNQCQELSEERFENCNTHRSRVGVC--KVTEXTFSA-----SNRLTGPPHPFKLTIRNPR 134
DB 74 BEVKCYSVLHATSYMDCPAIDATVFCGRDAVVAQPHDRV--QPFPEKGTLLRIVEPRV 131

QY 135 NSGSMFYVIVRLDDTKEPIDVFAIQLSVYQFANTATRGLYSKASCRTFGLPTVQLEAYL 194
DB 132 SDTGSYIIRVALAG-RNMSDIFRMAVIIRS-----SKWACHNSASSFOAHKCI 179

QY 195 RTEE--SWRNQAYVATEATTTSABATTPVATTSASLEAEHETFPWLENGVDHYEPT 252
DB 180 RYVDRMAFENYLIGHVGNLLDSDELHAIYNTPOSIS-TDINIITTFYDNSGTIYSP 238

QY 253 P---ANENSNVTVRLGT-MSPTLIGTVAAVY--SATIGLVIV---ISIVTRNMCPTHRK 303
DB 239 VFNLFNNSHVDAMNSTGMWNTVLKYLPRLIYFTSMIVLCIYALAIYIVCERCSPHR 298

QY 304 LDTVSODDEERQTRRESKFGPMVACEINKGADQDSSELVE 344
DB 299 ----TYIGEPKRSDE-----APLITSAVNESFQYDYNVKE 328

RESULT 15
Q91E39 PRELIMINARY; PRT; 369 AA.
AC Q91E39;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycoprotein 1.
GN US7.
OS Phocid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=47418;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB84;
RA Martina B., Osterhaus A.D.M.E., Harder T.C.;
RT "Identification and analysis of immunogenicity of the glycoprotein D
RT equivalent within the unique short segment of phocid herpesvirus-1.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ290955; CAC51466.1; -
DR InterPro: IPR002874; Herpes_gi.

```

```

DR InterPro: IPR007110; Ig-like.
DR Pfam: PF01688; Herpes_gi; 1.
DR SEQUENCE 369 AA; 42372 MW; 1587181704A7AA8A CRC64;

Query Match      7.5%; Score 140.5; DB 12; Length 369;
Best Local Similarity 25.1%; Pred. No. 0.00025;
Matches 57; Conservative 36; Mismatches 89; Indels 45; Gaps 10;

QY 22 GIVITGNHVSARIDDDHIVIVAPRP-EATIQQLFFMPGOR-PHKPYSGTVRVARFSDIT 79
DB 16 GIVIRGTYSMYNTSSGYTVYDDRDNFNTGVLFLDDQRLPVTNYSGHIEIYI---FN 72

QY 80 NQCQELSEERFENCNTHRSRVGVC--KVTEXTFSA-----SNRLTGPPHPFKLTIRNPR 133
DB 73 YSCVTYVQTIYVSCPRIHNNAFRSLIKVSKHQQLRINSSIEYG---VLEIKNPK 128

QY 134 PNDGSMFYVIVRLDDTKEPIDVFAIQLSVYQFANTATRGLYSKASCRTFGLPTVQLEAY 193
DB 129 PSDSGVYIFRVQLENNK--TDVFGISAFVYSPKSG--ENITKPDNSNOT----- 173

QY 194 LRTEESWRNQAAYVATEATTTSABA-----TTPTPVATTSASEL 232
DB 174 -----ENFTNHLVTPSTTISTKPFSETSHLNTFTPTDIPAPVCEV 213

Search completed: March 8, 2004, 02:44:33
Job time : 128 secs

```